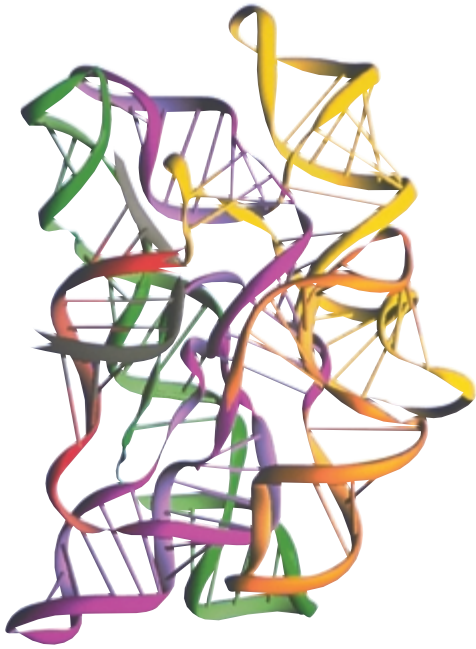
27 Nucleosides, Nucleotides, and Nucleic Acids



an RNA catalyst

In previous chapters, we studied two of the three major kinds of biopolymers—polysaccharides and proteins. Now we will look at the third—nucleic acids. There are two types of nucleic acids—**deoxyribonucleic acid (DNA)** and **ribonucleic acid (RNA)**. DNA encodes an organism's entire hereditary information and controls the growth and division of cells. In most organisms, the genetic information stored in DNA is transcribed into RNA. This information can then be translated for the synthesis of all the proteins needed for cellular structure and function.

DNA was first isolated in 1869 from the nuclei of white blood cells. Because this material was found in the nucleus and was acidic, it was called *nucleic acid*. Eventually, scientists found that the nuclei of all cells contain DNA, but it wasn't until 1944 that they realized that nucleic acids are the carriers of genetic information. In 1953, James Watson and Francis Crick described the three-dimensional structure of DNA—the famed double helix.

Studies that determined the structures of the nucleic acids and paved the way for the discovery of the DNA double helix were carried out by **Phoebus Levene** *and elaborated by* **Sir Alexander Todd**.

Phoebus Aaron Theodor Levene (1869–1940) *was born in Russia. When he immigrated to the United States with his family in 1891, his Russian name Fishel was changed to Phoebus. Because his medical school education had been interrupted, he returned to Russia to complete his studies. When he returned to the United States, he took chemistry courses at Columbia University. Deciding to forgo medicine for a career in chemistry, he went to Germany to study under Emil Fischer. He was a professor of chemistry at the Rockefeller Institute (now Rockefeller University).*

27.1 Nucleosides and Nucleotides

Nucleic acids are chains of five-membered-ring sugars linked by phosphate groups (Figure 27.1). The anomeric carbon of each sugar is bonded to a nitrogen of a heterocyclic compound in a β-glycosidic linkage. (Recall from Section 22.10 that a β-linkage is one in which the substituents at C-1 and C-4 are on the same side of the furanose ring.) Because the heterocyclic compounds are amines, they are commonly referred to as **bases**. In RNA the five-membered-ring sugar is D-ribose. In DNA it is 2-deoxy-D-ribose (D-ribose without an OH group in the 2-position).

Phosphoric acid links the sugars in both RNA and DNA. The acid has three dissociable OH groups with pK_a values of 1.9, 6.7, and 12.4. Each of the OH groups can react with an alcohol to form a *phosphomonoester*, a *phosphodiester*, or a *phosphotriester*. In nucleic acids the phosphate group is a **phosphodiester**.

THE STRUCTURE OF DNA: WATSON, CRICK, FRANKLIN, AND WILKINS

James D. Watson was born in Chicago in 1928. He graduated from the University of Chicago at the age of 19 and received a Ph.D. three years later from Indiana University. In 1951, as a postdoctoral fellow at Cambridge University, Watson worked on determining the three-dimensional structure of DNA.

Francis H. C. Crick was born in Northampton, England, in 1916. Originally trained as a physicist, Crick was involved in radar research during World War II. After the war, he entered Cambridge University to study for a Ph.D. in chemistry, which he received in 1953. He was a graduate student when he carried out his portion of the work that led to the proposal of the double helical structure of DNA.

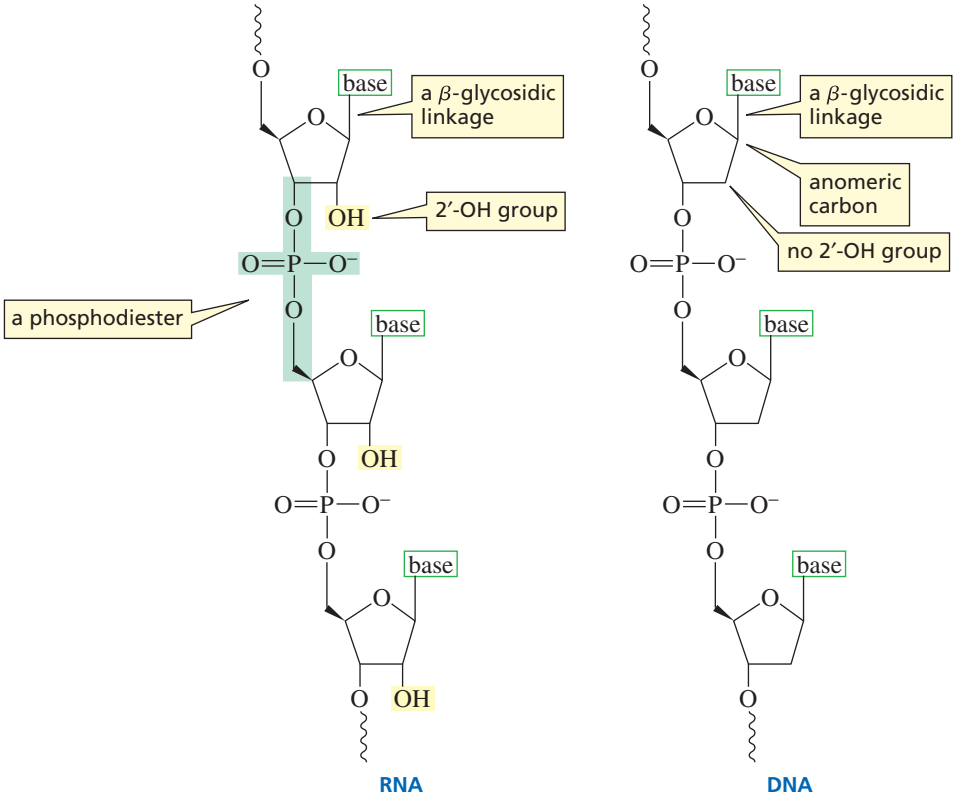
Rosalind Franklin was born in London in 1920. She graduated from Cambridge University and in 1942 quit her graduate studies to accept a position as a research officer in the British Coal Utilisation Research Association. After the war, she studied X-ray diffraction techniques in Paris. In 1951 she returned to England, accepting a position to develop an X-ray diffraction unit in the biophysics department at King's College. Her X-ray studies showed that DNA was a helix with phosphate groups on the outside of the molecule. Franklin died in 1958 without knowing the significance her work had played in determining the structure of DNA.

Watson and Crick shared the 1962 Nobel Prize in medicine or physiology with Maurice Wilkins for determining the double helical structure of DNA. Wilkins contributed X-ray studies that confirmed the double helical structure. Wilkins was born in New Zealand in 1916 and moved to England six years later with his parents. During World War II he joined other British scientists who were working with American scientists on the development of the atomic bomb.

Francis Crick and James Watson

Rosalind Franklin

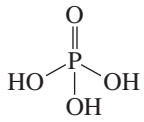
AU: Please identify by indicating left or right

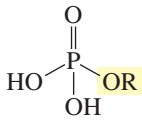


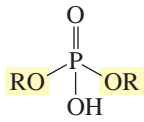
▲ **Figure 27.1**
Nucleic acids consist of a chain of five-membered-ring sugars linked by phosphate groups. Each sugar (D-ribose in RNA, 2′-deoxy-D-ribose in DNA) is bonded to a heterocyclic amine in a β-glycosidic linkage.

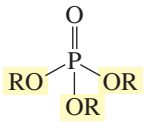
AU: ok prime here?

Alexander R. Todd (1907–1997) *was born in Scotland. He received two Ph.D. degrees, one from Johann Wolfgang Goethe University in Frankfurt (1931) and one from Oxford University (1933). He was a professor of chemistry at the University of Edinburgh, at the University of Manchester, and from 1944 to 1971 at Cambridge University. He was knighted in 1954 and was made a baron in 1962 (Baron Todd of Trumpington). For his work on nucleotides, he was awarded the 1957 Nobel Prize in chemistry.*

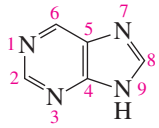

phosphoric acid

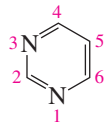

a phosphomonoester

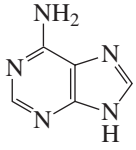

a phosphodiester

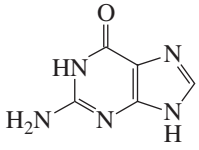

a phosphotriester

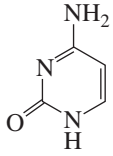
The vast differences in heredity among species and among members of the same species are determined by the sequence of the bases in DNA. Surprisingly, there are only four bases in DNA—two are substituted purines (adenine and guanine), and two are substituted pyrimidines (cytosine and thymine).

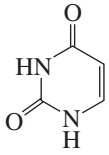

purine

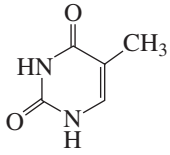

pyrimidine


adenine


guanine


cytosine


uracil

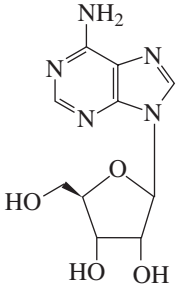

thymine

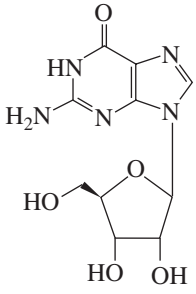
RNA also contains only four bases. Three (adenine, guanine, and cytosine) are the same as those in DNA, but the fourth base in RNA is uracil instead of thymine. Notice that thymine and uracil differ only by a methyl group—thymine is 5-methyluracil. The reason DNA contains thymine instead of uracil is explained in Section 27.14.

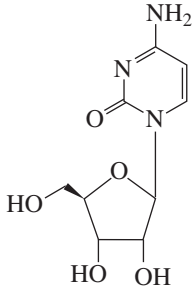
nucleoside = base + sugar

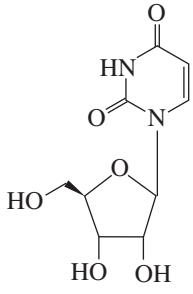
The purines and pyrimidines are bonded to the anomeric carbon of the furanose ring—purines at N-9 and pyrimidines at N-1—in a β-glycosidic linkage. A compound containing a base bonded to D-ribose or to 2-deoxy-D-ribose is called a **nucleoside**. In a nucleoside the ring positions of the sugar are indicated by primed numbers to distinguish them from the ring positions of the base. This is why the sugar component of DNA is referred to as 2′-deoxy-D-ribose. Notice the difference in the base names and their corresponding nucleoside names in Table 27.1. For example, adenine is the base, whereas adenosine is the nucleoside. Similarly, cytosine is the base, whereas cytidine is the nucleoside, and so forth. Because uracil is found only in RNA, it is shown attached to D-ribose but not to 2-deoxy-D-ribose; because thymine is found only in DNA, it is shown attached to 2-deoxy-D-ribose but not to D-ribose.

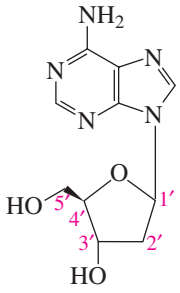
nucleosides

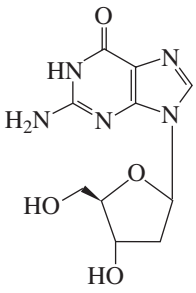

adenosine

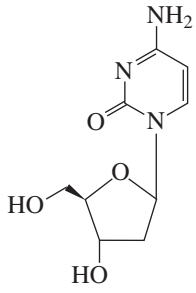

guanosine


cytidine


uridine


2′-deoxyadenosine


2′-deoxyguanosine


2′-deoxycytidine

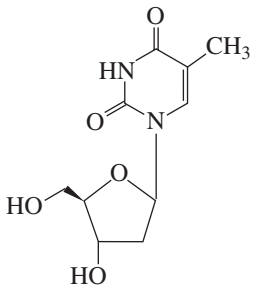

thymidine

Table 27.1 The Names of the Bases, the Nucleosides, and the Nucleotides

Base	Nucleoside		Ribonucleotide	Deoxyribonucleotide
Adenine	Adenosine	2′-Deoxyadenosine	Adenosine 5′-phosphate	2′-Deoxyadenosine 5′-phosphate
Guanine	Guanosine	2′-Deoxyguanosine	Guanosine 5′-phosphate	2′-Deoxyguanosine 5′-phosphate
Cytosine	Cytidine	2′-Deoxycytidine	Cytidine 5′-phosphate	2′-Deoxycytidine 5′-phosphate
Thymine	—	Thymidine	—	Thymidine 5′-phosphate
Uracil	Uridine	—	Uridine 5′-phosphate	

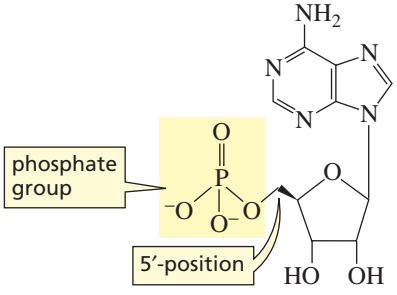
PROBLEM 1

In acidic solutions, nucleosides are hydrolyzed to a sugar and a heterocyclic base. Propose a mechanism for this reaction.

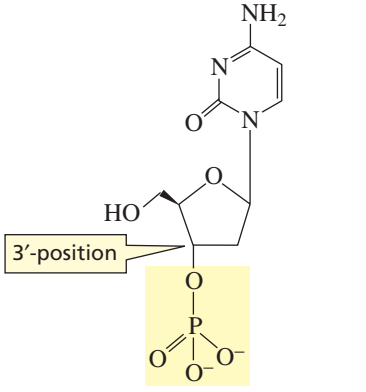
A **nucleotide** is a nucleoside with either the 5′- or the 3′-OH group bonded in an ester linkage to phosphoric acid. The nucleotides of RNA—where the sugar is D-ribose—are more precisely called **ribonucleotides**, whereas the nucleotides of DNA—where the sugar is 2-deoxy-D-ribose—are called **deoxyribonucleotides**.

nucleotide = base + sugar + phosphate

nucleotides



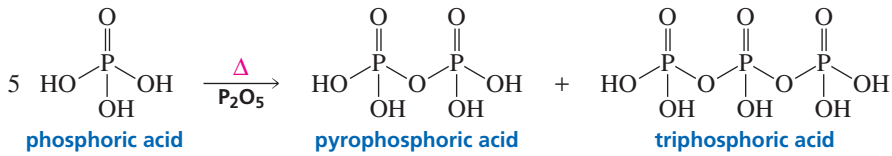
adenosine 5′-monophosphate
a ribonucleotide



2′-deoxycytidine 3′-monophosphate
a deoxyribonucleotide

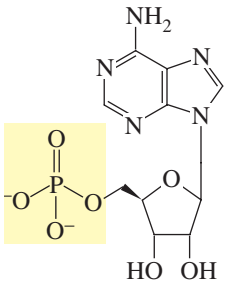

3-D Molecules: Bases; Nucleosides; Nucleotides

When phosphoric acid is heated with P_2O_5 it loses water, forming a phosphoanhydride called pyrophosphoric acid. Its name comes from *pyr*, the Greek word for "fire." Thus, pyrophosphoric acid is prepared by "fire"—that is, by heating. Triphosphoric acid and higher polyphosphoric acids are also formed.

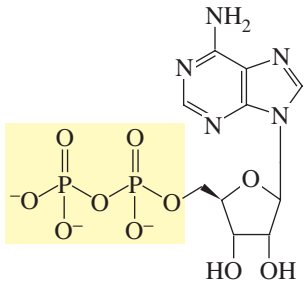


phosphoric acid — pyrophosphoric acid — triphosphoric acid

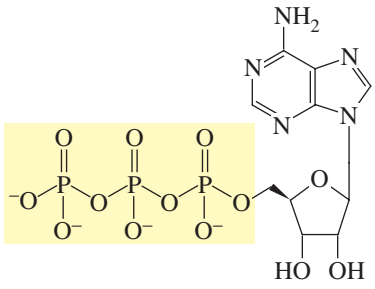
Because phosphoric acid can form an anhydride, nucleotides can exist as monophosphates, diphosphates, and triphosphates. They are named by adding *monophosphate* or *diphosphate* or *triphosphate* to the name of the nucleoside.



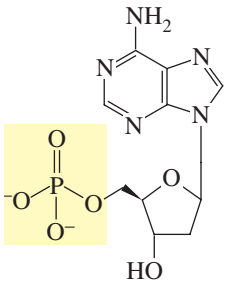
adenosine 5′-monophosphate
AMP



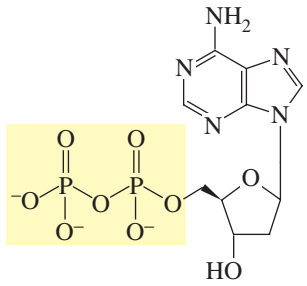
adenosine 5′-diphosphate
ADP



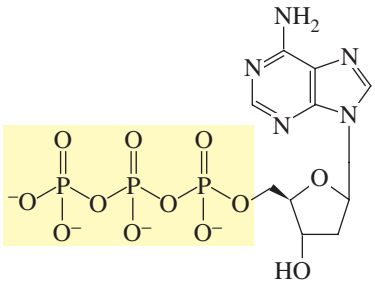
adenosine 5′-triphosphate
ATP



2′-deoxyadenosine 5′-monophosphate
dAMP



2′-deoxyadenosine 5′-diphosphate
dADP



2′-deoxyadenosine 5′-triphosphate
dATP

PROBLEM 2

Draw the structure for each of the following:

a. dCDP
b. dTTP
c. dUMP
d. UDP
e. guanosine 5′-triphosphate
f. adenosine 3′-monophosphate

27.2 ATP: The Carrier of Chemical Energy

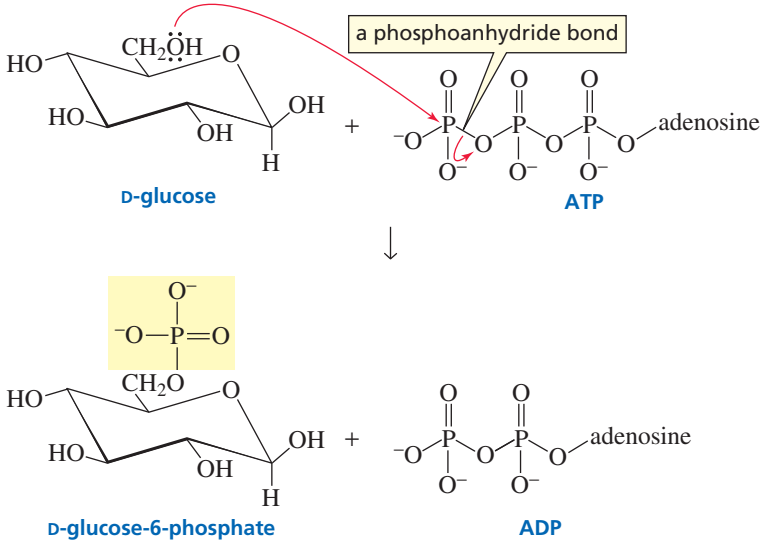
All cells require energy to ensure their survival and reproduction. They get the energy they need by converting nutrients into a chemically useful form of energy. The most important form of chemical energy is adenosine 5′-triphosphate (ATP). The importance of ATP to biological reactions is shown by its turnover rate in humans—each day, a person uses an amount of ATP equivalent to his or her body weight. ATP is known as the universal carrier of chemical energy because, as it is commonly stated, "the energy of hydrolysis of ATP converts endergonic reactions into exergonic reactions."

In other words, the ability of ATP to enable otherwise unfavorable reactions to occur is attributed to the large amount of energy released when ATP is hydrolyzed, which can be used to drive an endergonic reaction. For example, the reaction of

D-glucose with hydrogen phosphate to form D-glucose-6-phosphate is endergonic ($\Delta G^{\circ\prime}$ = +3.3 kcal/mol or +13.8 kJ/mol).[1] The hydrolysis of ATP, on the other hand, is highly exergonic ($\Delta G^{\circ\prime}$ = −7.3 kcal/mol or −30.5 kJ/mol). When the two reactions are added together (the species occurring on both sides of the reaction arrow cancel), the net reaction is exergonic ($\Delta G^{\circ\prime}$ = −4.0 kcal/mol or −16.7 kJ/mol). Thus, the energy released from the hydrolysis of ATP is more than enough to drive the phosphorylation of D-glucose. Two reactions in which the energy of one is used to drive the other are known as *coupled reactions*.

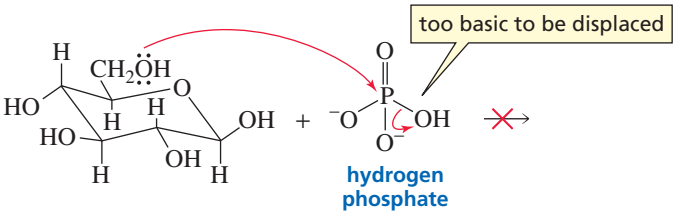
	$\Delta G^{\circ\prime}$		
D-glucose + hydrogen phosphate ⟶ D-glucose-6-phosphate + H_2O	+ 3.3 kcal/mol	or	+ 13.8 kJ/mol
ATP + H_2O ⟶ ADP + hydrogen phosphate	− 7.3 kcal/mol	or	− 30.5 kJ/mol
D-glucose + ATP ⟶ D-glucose-6-phosphate + ADP	− 4.0 kcal/mol	or	− 16.7 kJ/mol

This nonmechanistic description of ATP's power makes ATP look like a magical source of energy. Let's look at the mechanism of the reaction to see what really happens. The reaction is a simple one-step nucleophilic substitution reaction. The 6-OH group of glucose attacks the terminal phosphate of ATP, breaking a **phosphoanhydride bond** without forming an intermediate. Essentially it is an S_N2 reaction with an adenosine pyrophosphate leaving group.




3-D Molecules: Adenosine 5′-triphosphate (ATP)

Now we have a chemical understanding of why the phosphorylation of glucose requires ATP. Without ATP, the 6-OH group of D-glucose would have to displace a very basic ^{-}OH group from hydrogen phosphate. With ATP, the 6-OH group of D-glucose displaces the weakly basic ADP.



[1]The prime in $\Delta G^{\circ\prime}$ indicates that two additional parameters have been added to the ΔG° defined in Section 3.7—the reaction occurs in aqueous solution at pH = 7 and the concentration of water is assumed to be constant.

Although the phosphorylation of glucose is described as being driven by the "hydrolysis" of ATP, you can see from the mechanism that glucose does not react with hydrogen phosphate and that ATP is not hydrolyzed because it does not react with water. In other words, neither of the coupled reactions actually occurs. Instead, the phosphate group of ATP is transferred directly to D-glucose.

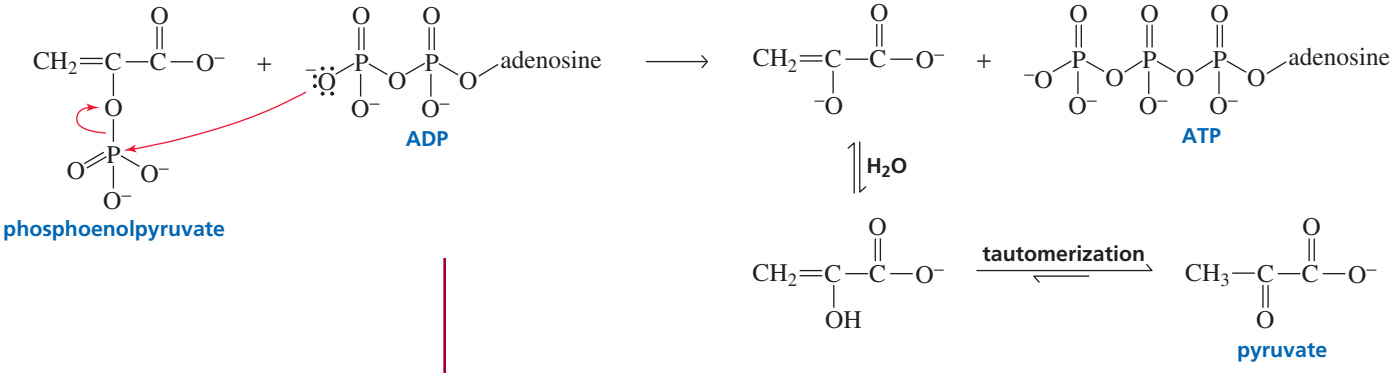
The transfer of a phosphate group from ATP to D-glucose is an example of a **phosphoryl transfer reaction**. There are many phosphoryl transfer reactions in biological systems. In all of these reactions, the electrophilic phosphate group is transferred to a nucleophile as a result of breaking a phosphoanhydride bond. This example of a phosphoryl transfer reaction demonstrates the actual chemical function of ATP—*it provides a reaction pathway involving a good leaving group for a reaction that cannot occur (or would occur very slowly) because of a poor leaving group.*

ATP provides a reaction pathway with a good leaving group for a reaction that cannot occur because of a poor leaving group.

PROBLEM 3 SOLVED

The hydrolysis of phosphoenolpyruvate is so highly exergonic ($\Delta G^{\circ\prime} = -14.8$ kcal/mol or -61.9 kJ/mol) that it can be used to "drive the formation" of ATP from ADP and hydrogen phosphate ($\Delta G^{\circ\prime} = +7.3$ kcal/mol or $+30.5$ kJ/mol). Propose a mechanism for this reaction.

SOLUTION As we saw in the example with ATP, neither of the coupled reactions actually occurs: phosphoenolpyruvate does not react with water and ADP does not react with hydrogen phosphate. Just as ATP "drives the formation" of glucose-6-phosphate by supplying glucose (a nucleophile) with a phosphate that has a good leaving group (ADP), phosphoenolpyruvate "drives the formation" of ATP by supplying ADP (a nucleophile) with a phosphate that has a good leaving group (pyruvate).



PROBLEM 4◆

Why is pyruvate a good leaving group?

PROBLEM 5◆

Several important biomolecules and the $\Delta G^{\circ\prime}$ values for their hydrolysis are listed here. Which of them "hydrolyzes" with sufficient energy to "drive the formation" of ATP?

glycerol-1-phosphate:
-2.2 kcal/mol (-9.2 kJ/mol)

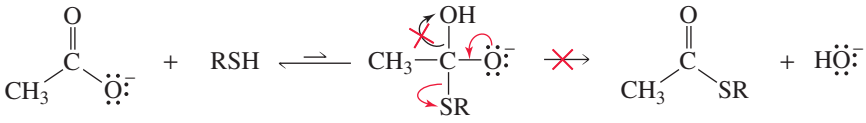
phosphocreatine:
-11.8 kcal/mol (-49.4 kJ/mol)

fructose-6-phosphate:
-3.8 kcal/mol3 (-15.9 kJ/mol)

glucose-6-phosphate:
-3.3 kcal/mol (13.8 kJ/mol).

27.3 Three Mechanisms for Phosphoryl Transfer Reactions

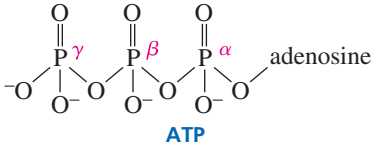
There are three possible mechanisms for a phosphoryl transfer reaction. We will illustrate them using the following nucleophilic acyl substitution reaction.



This reaction does not occur without ATP because the negatively charged carboxylate ion resists nucleophilic attack and, if the tetrahedral intermediate could be formed, the incoming nucleophile is a weaker base than the base that would have to be expelled from the tetrahedral intermediate to form the thioester. In other words, the thiol would be expelled from the tetrahedral intermediate, reforming the carboxylate ion (Section 17.5).

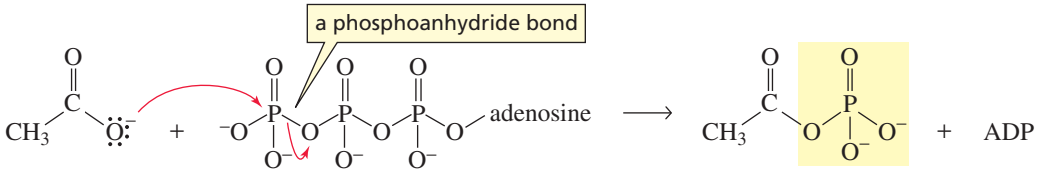
If ATP is added to the reaction mixture, the reaction occurs. The carboxylate ion attacks one of the phosphate groups of ATP, breaking a phosphoanhydride bond. This puts a leaving group on the carboxyl group that can be displaced by the thiol. There are three possible mechanisms for the reaction of a nucleophile with ATP because each of the three phosphorus atoms of ATP can undergo nucleophilic attack. Each mechanism puts a different phosphate leaving group on the nucleophile.

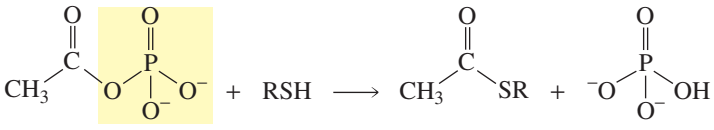
Tutorial: Phosphoryl transfer reactions



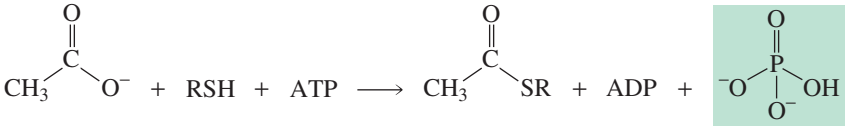
If the carboxylate ion attacks the γ-phosphorus of ATP, an **acyl phosphate** is formed. The acyl phosphate then reacts with the thiol in a nucleophilic acyl substitution reaction (Section 17.5) to form the thioester.

nucleophilic attack on the γ-phosphorus



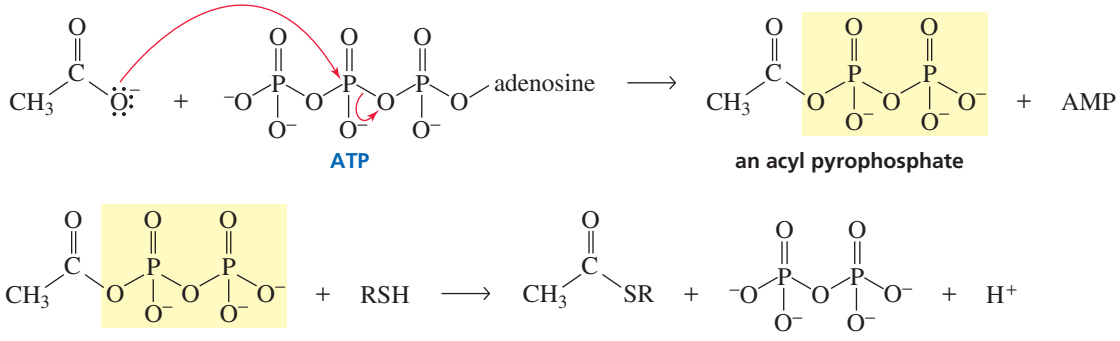


overall reaction

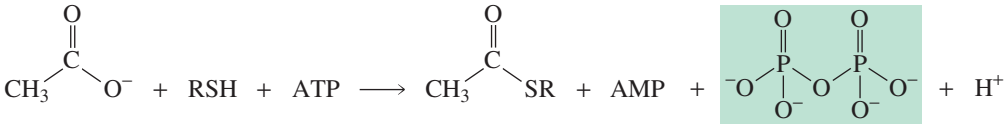


If the carboxylate ion attacks the β-phosphorus of ATP, an **acyl pyrophosphate** is formed. The acyl phosphate then reacts with the thiol in a nucleophilic acyl substitution reaction to form the thioester.

nucleophilic attack on the β-phosphorus

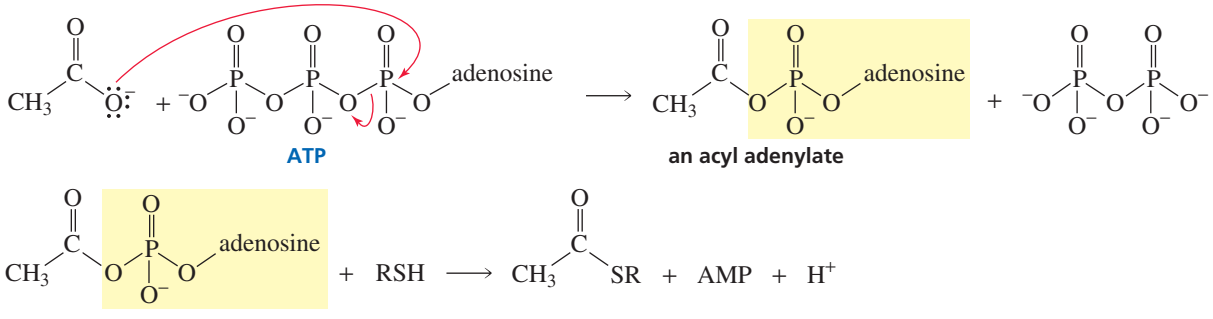


overall reaction

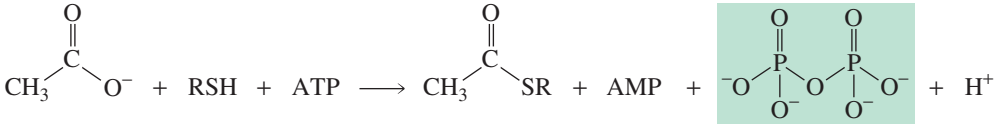


In the third possible mechanism, the carboxylate ion attacks the α-phosphorus of ATP, forming an **acyl adenylate**. The acyl phosphate then reacts with the thiol in a nucleophilic acyl substitution reaction to form the thioester.

nucleophilic attack on the α-phosphorus



overall reaction

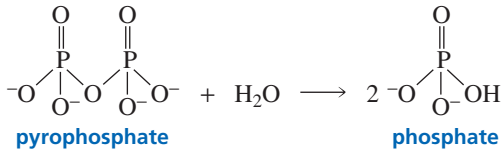


In Section 17.20 we saw that carboxylic acids in biological systems can be activated by being converted into acyl phosphates, acyl pyrophosphates, and acyl adenylates. Each of these three mechanisms puts a leaving group on the carboxylic acid that can easily be displaced by a nucleophile. The only difference in the three mechanisms is the particular phosphate atom that is attacked by the nucleophile and the nature of the intermediate that is formed.

Many different nucleophiles react with ATP in biological systems. Whether nucleophilic attack occurs on the α-, β-, or γ-phosphorus in any particular reaction depends

on the enzyme catalyzing the reaction (Section 27.5). Mechanisms involving nucleophilic attack on the γ-phosphorus form ADP and phosphate as side products, whereas mechanisms involving nucleophilic attack on the α- or β-phosphorus form AMP and pyrophosphate as side products.

When pyrophosphate is one of the side products, it is subsequently hydrolyzed to two equivalents of phosphate. Consequently, in reactions in which pyrophosphate is formed as a product, its subsequent hydrolysis drives the reaction to the right, ensuring its irreversibility.



Therefore, enzyme-catalyzed reactions in which irreversibility is important take place by one of the mechanisms that form pyrophosphate as a product (attack on the α- or β-phosphorus of ATP). For example, both the reaction that links nucleotide subunits to form nucleic acids (Section 27.7) and the reaction that binds an amino acid to a tRNA (the first step in translating RNA into a protein; Section 27.12) involve nucleophilic attack on the α-phosphorus of ATP.

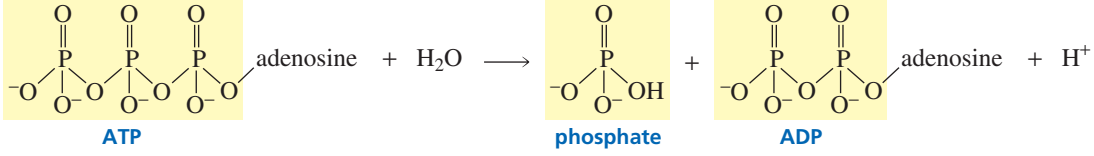
PROBLEM 6

The β-phosphorus of ATP has two phosphoanhydride linkages, but only the one linking the β-phosphorus to the α-phosphorus is broken in phosphoryl transfer reactions. Explain why the one linking the β-phosphorus to the γ-phosphorus is never broken.

27.4 The "High-Energy" Character of Phosphoanhydride Bonds

Because the hydrolysis of a phosphoanhydride bond is a highly exergonic reaction, phosphoanhydride bonds are called **"high-energy bonds."** The term "high-energy" in this context means that a lot of energy is released when a reaction occurs that causes the bond to break. Do not confuse it with "bond energy," the term chemists use to describe how difficult it is to break a bond. A bond with a *high bond energy* is hard to break, whereas a *high-energy bond* breaks readily.

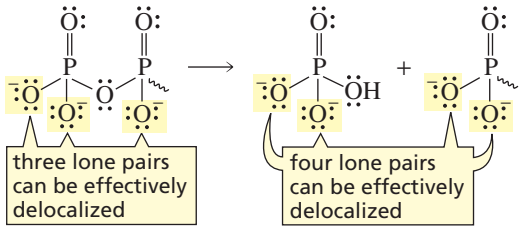
Why is the hydrolysis of a phosphoanhydride bond so exergonic? In other words, why is the $\Delta G^{\circ\prime}$ value for its hydrolysis large and negative? A large negative $\Delta G^{\circ\prime}$ means that the products of the reaction are much more stable than the reactants. Let's look at ATP and its hydrolysis products, phosphate and ADP, to see why this is so.



Three factors contribute to the greater stability of ADP and phosphate compared to ATP:

1. **Greater electrostatic repulsion in ATP.** At physiological pH (pH = 7.3), ATP has 3.3 negative charges, ADP has 2.8 negative charges, and phosphate has 1.1 negative charges (Section 1.20). Because of ATP's greater negative charge, more electrostatic repulsions are present in ATP than in ADP or phosphate. Electrostatic repulsions destabilize a molecule.

2. **More solvation in the products.** Negatively charged ions are stabilized in an aqueous solution by solvation. Because the reactant has 3.3 negative charges, while the sum of the negative charges on the products is 3.9 (2.8 + 1.1), there is more solvation in the products than in the reactant.
3. **Greater resonance stabilization in the products.** Delocalization of a lone pair on the oxygen joining the two phosphorus atoms is not very effective because it puts a positive charge on an oxygen that is next to a partially positively charged phosphorus atom. When the phosphoanhydride bond breaks, one additional lone pair can be effectively delocalized.



Similar factors explain the large negative $\Delta G^{\circ\prime}$ when ATP is hydrolyzed to AMP and pyrophosphate, and when pyrophosphate is hydrolyzed to two equivalents of phosphate.

PROBLEM 7 SOLVED

ATP has pK_a values of 0.9, 1.5, 2.3, and 7.7; ADP has pK_a values of 0.9, 2.8, and 6.8; and phosphoric acid has pK_a values of 1.9, 6.7, and 12.4.
Do the calculation showing that at pH 7.3:

a. the charge on ATP is −3.3

b. the charge on ADP is −2.8

c. the charge on phosphate is −1.1

SOLUTION TO 7a Because pH 7.3 is much more basic than the pK_a values of the first three ionizations of ATP, we know that these three groups will be entirely in their basic forms at that pH, giving ATP three negative charges. We need to determine what fraction of the group with pK_a 7.7 will be in its basic form at pH 7.3.

$$\frac{\text{concentration in the basic form}}{\text{total concentration}} = \frac{[A^-]}{[A^-] + [HA]}$$

$$[A^-] = \text{concentration of the basic form}$$

$$[HA] = \text{concentration of the acidic form}$$

Because this equation has two unknowns, one of the unknowns must be expressed in terms of the other unknown. Using the definition of the acid dissociation constant (K_a), we can define $[HA]$ in terms of $[A^-]$, K_a, and $[H^+]$.

$$K_a = \frac{[A^-][H^+]}{[HA]}$$

$$[HA] = \frac{[A^-][H^+]}{K_a}$$

$$\frac{[A^-]}{[A^-] + [HA]} = \frac{[A^-]}{[A^-] + \dfrac{[A^-][H^+]}{K_a}} = \frac{K_a}{K_a + [H^+]}$$

Now we can calculate the fraction of the group with pK_a 7.7 that will be in the basic form.

$$\frac{K_a}{K_a + [H^+]} = \frac{2.0 \times 10^{-8}}{2.0 \times 10^{-8} + 5.0 \times 10^{-8}} = 0.3$$

total negative charge on ATP = 3.0 + 0.3 = 3.3

27.5 Kinetic Stability of ATP in the Cell

Although ATP reacts readily in enzyme-catalyzed reactions, it reacts quite slowly in the absence of an enzyme. For example, carboxylic acid anhydrides hydrolyze in a matter of minutes, but ATP takes several weeks to hydrolyze. The low rate of ATP hydrolysis is important because it allows ATP to exist in the cell until it is needed for an enzyme-catalyzed reaction.

The negative charges on ATP are what make it relatively unreactive. These negative charges repel the approach of nucleophiles. When ATP is bound at an active site of an enzyme, it complexes with magnesium (Mg^{2+}), which decreases the overall negative charge on ATP. (This is why ATP-requiring enzymes also require metal ions; Section 25.5.) The other two negative charges can be stabilized by positively charged groups such as arginine or lysine residues at the active site, as shown in Figure 27.2. In this form, ATP is readily approached by nucleophiles, so ATP reacts rapidly in an enzyme-catalyzed reaction, but only very slowly in the absence of the enzyme.

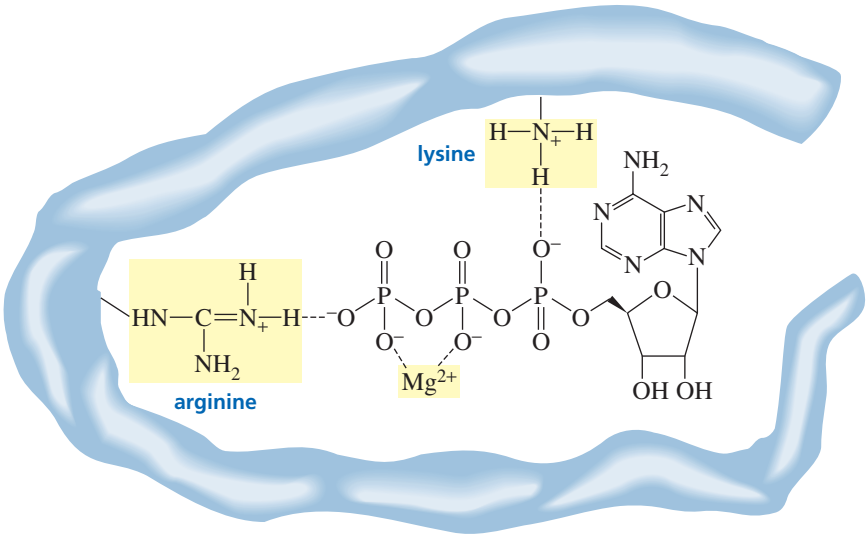
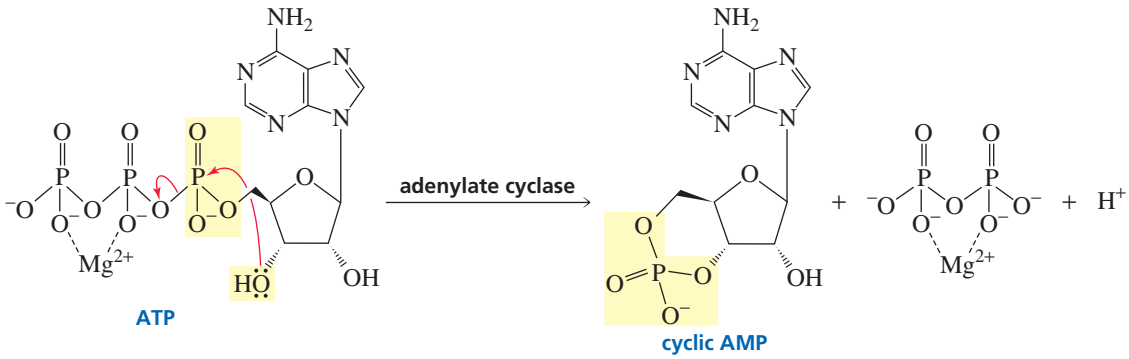


Figure 27.2
The interactions between ATP, Mg^{2+}, and arginine and lysine residues at the active site of an enzyme.

27.6 Other Important Nucleotides

ATP is not the only biologically important nucleotide. Guanosine 5′-triphosphate (GTP) is used in place of ATP in some phosphoryl transfer reactions. We have also seen in Sections 25.2 and 25.3 that dinucleotides are used as oxidizing agents (NAD^+, $NADP^+$, FAD, FMN) and reducing agents (NADH, NADPH, $FADH_2$, $FMNH_2$).

Another important nucleotide is adenosine 3′,5′-monophosphate, commonly known as cyclic AMP. Cyclic AMP is called a "second messenger" because it serves as a link between several hormones (the first messengers) and certain enzymes that regulate cellular function. Secretion of certain hormones, such as adrenaline, activates adenylate cyclase, the enzyme responsible for the synthesis of cyclic AMP from ATP. Cyclic AMP then activates an enzyme, generally by phosphorylating it. Cyclic nucleotides are so important in regulating cellular reactions that an entire scientific journal is devoted to these processes.

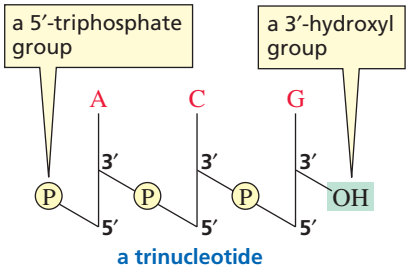


3-D Molecule: Cyclic AMP

PROBLEM 8

What products would be obtained from the hydrolysis of cyclic AMP?

27.7 The Nucleic Acids



Nucleic acids are composed of long strands of nucleotide subunits linked by phosphodiester bonds. These linkages join the 3′-OH group of one nucleotide to the 5′-OH group of the next nucleotide (Figure 27.1). A **dinucleotide** contains two nucleotide subunits, an **oligonucleotide** contains three to ten subunits, and a **polynucleotide** contains many subunits. DNA and RNA are polynucleotides. Notice that the nucleotide at one end of the strand has an unlinked 5′-triphosphate group, and the nucleotide at the other end of the strand has an unlinked 3′-hydroxyl group.

Nucleotide triphosphates are the starting materials for the biosynthesis of nucleic acids. DNA is synthesized by enzymes called *DNA polymerases*, and RNA is synthesized by enzymes called *RNA polymerases*. The nucleotide strand is formed as a result of nucleophilic attack by a 3′-OH group of one nucleotide triphosphate on the α-phosphorus of another nucleotide triphosphate, breaking a phosphoanhydride bond and eliminating pyrophosphate (Figure 27.3). This means that the growing polymer is synthesized in the 5′ ⟶ 3′ direction; in other words, new nucleotides are added to the 3′-end. Pyrophosphate is subsequently hydrolyzed, which makes the reaction irreversible (Section 27.3). RNA strands are biosynthesized in the same way, using ribonucleotides instead of 2′-deoxyribonucleotides. The **primary structure** of a nucleic acid is the sequence of bases in the strand.

Watson and Crick concluded that DNA consists of two strands of nucleic acids with the sugar–phosphate backbone on the outside and the bases on the inside. The chains are held together by hydrogen bonds between the bases on one strand and the bases on the other strand (Figure 27.4). The width of the double-stranded molecule is relatively constant, so a purine must pair with a pyrimidine. If the larger purines paired, the strand would bulge; if the smaller pyrimidines paired, the strands would have to contract to bring the two pyrimidines close enough to form hydrogen bonds.

Critical to Watson and Crick's proposal for the secondary structure of DNA were experiments carried out by Erwin Chargaff. These experiments showed that the number of adenines in DNA equals the number of thymines and the number of guanines equals the number of cytosines. Chargaff also noted that the number of adenines and thymines relative to the number of guanines and cytosines is characteristic of a given species but varies from species to species. In human DNA, for example, 60.4% of the bases are adenines and thymines, whereas 74.2% of them are adenines and thymines in the DNA of the bacterium *Sarcina lutea*.

Chargaff's data showing that [adenine] = [thymine] and [guanine] = [cytosine] could be explained if adenine (A) always paired with thymine (T) and guanine (G) always paired with cytosine (C). This means the two strands are *complementary*—where there is an A in one strand, there is a T in the opposing strand, and where there is a G in

Erwin Chargaff *was born in Austria in 1905 and received a Ph.D. from the University of Vienna. To escape Hitler, he came to the United States in 1935, becoming a professor at Columbia University College of Physicians and Surgeons. He modified paper chromatography, a technique developed to identify amino acids (Section 23.5), so that it could be used to quantify the different bases in a sample of DNA.*

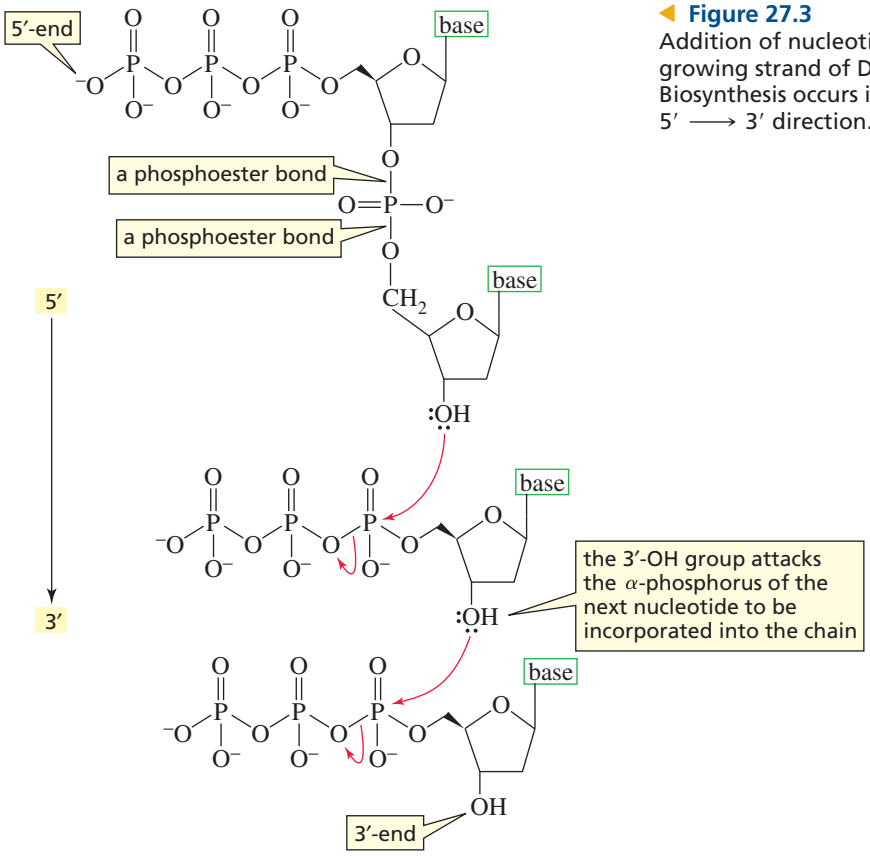


Figure 27.3
Addition of nucleotides to a growing strand of DNA. Biosynthesis occurs in the 5′ ⟶ 3′ direction.

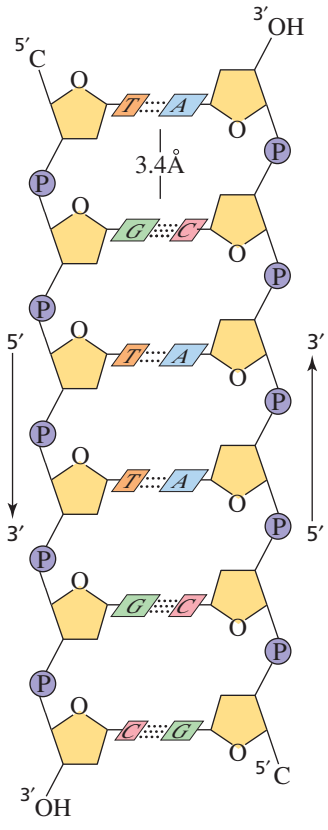


Figure 27.4
Complementary base pairing in DNA. Adenine (a purine) always pairs with thymine (a pyrimidine); guanine (a purine) always pairs with cytosine (a pyrimidine).

[A] = [T]
[G] = [C]

one strand there is a C in the other strand (Figure 27.4). Thus, if you know the sequence of bases in one strand, you can figure out the sequence of bases in the other strand.

What causes adenine to pair with thymine rather than with cytosine (the other pyrimidine)? The base pairing is dictated by hydrogen bonding. Learning that the bases exist in the keto form (Section 19.2) allowed Watson to explain the pairing.[2] Adenine forms two hydrogen bonds with thymine but would form only one hydrogen bond with cytosine. Guanine forms three hydrogen bonds with cytosine but would form only one hydrogen bond with thymine (Figure 27.5). The N—H----N and

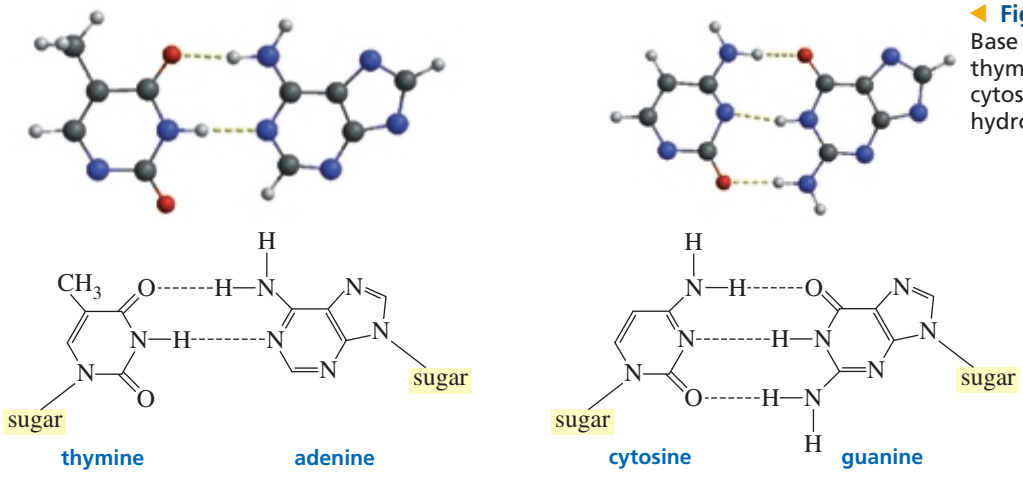
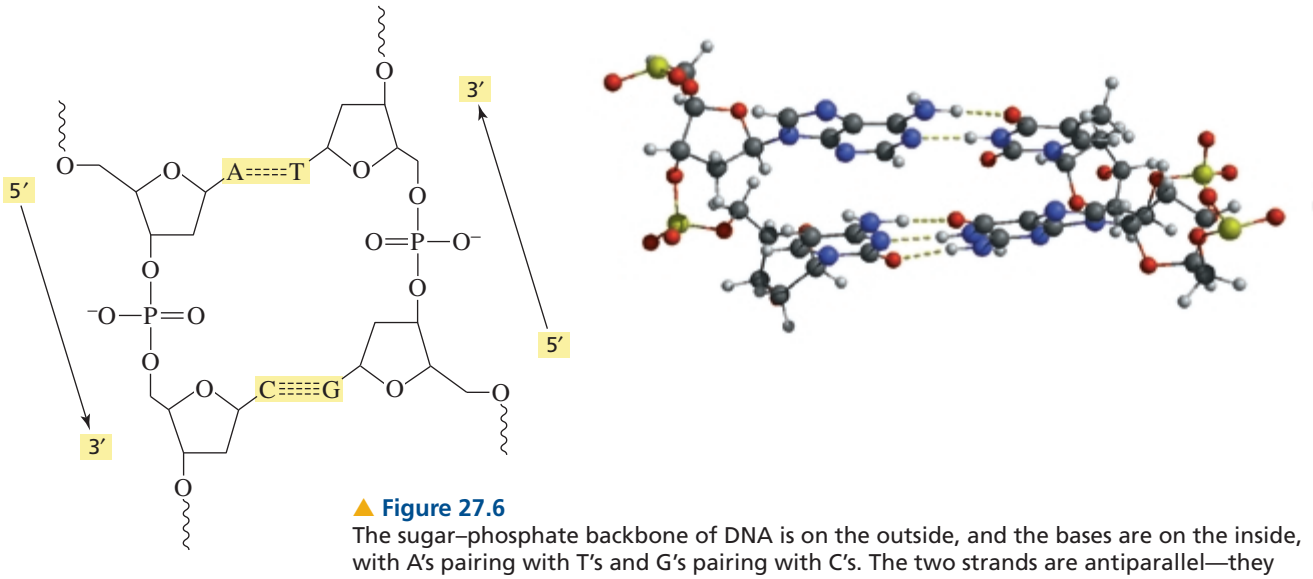


Figure 27.5
Base pairing in DNA: Adenine and thymine form two hydrogen bonds; cytosine and guanine form three hydrogen bonds.

[2]Watson was having difficulty understanding the base pairing in DNA because he thought the bases existed in the enol form (see Problem 10). When Jerry Donohue, an American crystallographer, informed him that the bases more likely existed in the keto form, Chargaff's data could easily be explained by hydrogen bonding between adenine and thymine and between guanine and cytosine.



▲ **Figure 27.6**
The sugar–phosphate backbone of DNA is on the outside, and the bases are on the inside, with A's pairing with T's and G's pairing with C's. The two strands are antiparallel—they run in opposite directions.

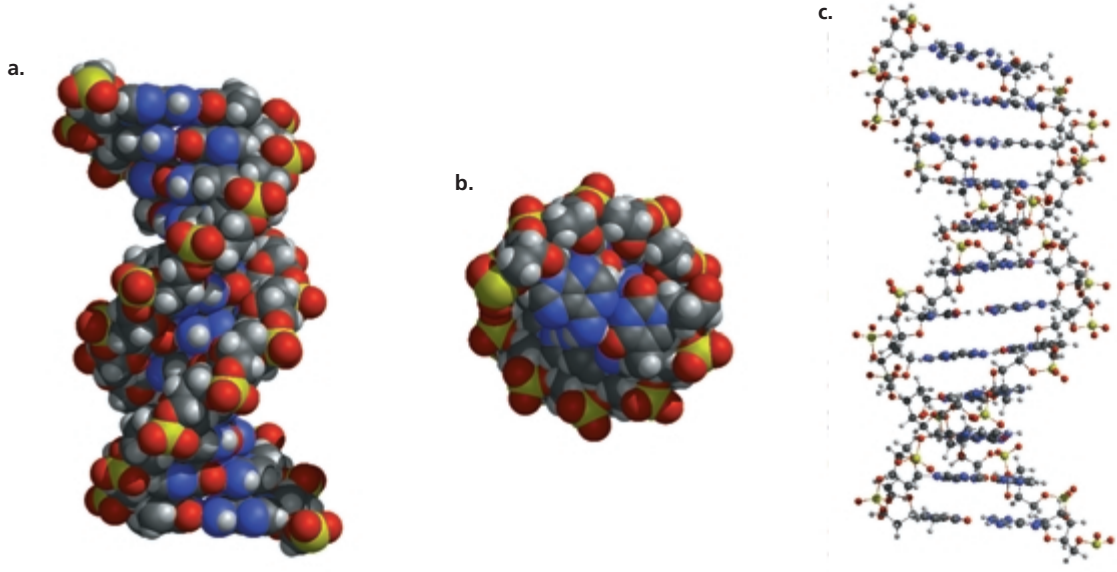
N—H····O bonds that hold the bases together are all about the same length (2.9 ± 0.1 Å).

The two DNA strands are antiparallel—they run in opposite directions, with the sugar–phosphate backbone on the outside and the bases on the inside (Figures 27.4 and 27.6). By convention, the sequence of bases in a polynucleotide is written in the 5′ ⟶ 3′ direction (the 5′-end is on the left).

ATGAGCCATGTAGCCTAATCGGC

5′-end 3′-end

The DNA strands are not linear but are twisted into a helix around a common axis (see Figure 27.7a). The base pairs are planar and parallel to each other on the inside of the helix (Figures 27.7b and c). The secondary structure is therefore known as a



▲ **Figure 27.7**
(a) The DNA double helix. (b) View looking down the long axis of the helix. (c) The bases are planar and parallel on the inside of the helix.

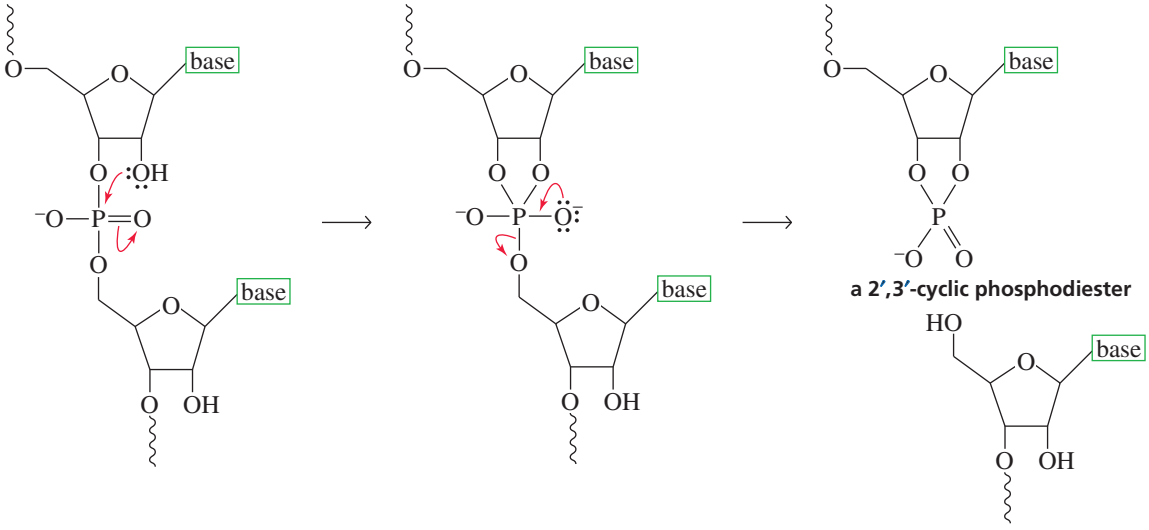
double helix. The double helix resembles a ladder (the base pairs are the rungs) twisted around an axis running down through its rungs (Figures 27.4 and 27.7c). The sugar–phosphate backbone is wrapped around the bases. The phosphate OH group has a pK_a of about 2, so it is in its basic form (negatively charged) at physiological pH. The negatively charged backbone repels nucleophiles, thereby preventing cleavage of the phosphodiester bonds.

Unlike DNA, RNA is easily cleaved because the 2′-OH group of ribose can act as the nucleophile that cleaves the strand (Figure 27.8). This explains why the 2′-OH group is absent in DNA. To preserve the genetic information, DNA must remain intact throughout the life span of a cell. Cleavage of DNA would have disastrous consequences for the cell and for life itself. RNA, in contrast, is synthesized as it is needed and is degraded once it has served its purpose.

Hydrogen bonding between base pairs is just one of the forces holding the two strands of the DNA double helix together. The bases are planar aromatic molecules that stack on top of one another. Each pair is slightly rotated with respect to the next pair, like a partially spread-out hand of cards. There are favorable van der Waals interactions between the mutually induced dipoles of adjacent pairs of bases. These interactions, known as **stacking interactions**, are weak attractive forces, but when added together they contribute significantly to the stability of the double helix. Stacking interactions are strongest between two purines and weakest between two pyrimidines. Confining the bases to the inside of the helix has an additional stabilizing effect—it reduces the surface area of the relatively nonpolar residues exposed to water. This increases the entropy of the surrounding water molecules (Section 23.14).

3-D Molecules:
Adenine–thymine base pair;
Guanine–cytosine base pair

3-D Molecule:
DNA double helix



▲ **Figure 27.8**
Hydrolysis of RNA. The 2′-OH group acts as an intramolecular nucleophilic catalyst. It has been estimated that RNA is hydrolyzed 3 billion times faster than DNA.

PROBLEM 9

Indicate whether each functional group of the five heterocyclic bases in nucleic acids can function as a hydrogen bond acceptor (A), a hydrogen bond donor (D), or both (D/A).

PROBLEM 10

Using the D, A, and D/A designations in Problem 9, explain how base pairing would be affected if the bases existed in the enol form.

PROBLEM 11

The 2′,3′-cyclic phosphodiester, which is formed when RNA is hydrolyzed (Figure 27.8), reacts with water, forming a mixture of nucleotide 2′- and 3′-phosphates. Propose a mechanism for this reaction.

PROBLEM 12◆

If one of the strands of DNA has the following sequence of bases running in the 5′ ⟶ 3′ direction,

5′—G—G—A—C—A—A—T—C—T—G—C—3′

a. What is the sequence of bases in the complementary strand?

b. What base is closest to the 5′-end in the complementary strand?

27.8 Helical Forms of DNA

Naturally occurring DNA can exist in the three different helical forms shown in Figure 27.9. The B- and A-helices are both right-handed. The B-helix is the predominant form in aqueous solution, while the A-helix is the predominant form in nonpolar solvents. Nearly all the DNA in living organisms is in a B-helix. The Z-helix is a left-handed helix. It occurs in regions where there is a high content of G—C base pairs. The A-helix is shorter (for a given number of base pairs) and about 3% broader than the B-helix, which is shorter and broader than the Z-helix.

Helices are characterized by the number of bases per 360° turn and the distance (the rise) between adjacent base pairs. A-DNA has 11 base pairs per turn and a 2.3 Å rise; B-DNA has 10 base pairs per turn and a 3.4 Å rise; and Z-DNA has 12 base pairs per turn and a 3.8 Å rise.

If you examine Figure 27.9, you will see that there are two kinds of alternating grooves in DNA. In B-DNA the **major groove** is wider than the **minor groove**. Cross sections of the double helix show that one side of each base pair faces into the major groove and the other side faces into the minor groove (Figure 27.10).

Proteins and other molecules can bind to the grooves. The hydrogen-bonding properties of the functional groups facing into each groove determine what kind of molecules will bind to the groove. Mitomycin is a naturally occurring compound that has been found to have both antibacterial activity and anticancer activity. It works by binding to the minor groove of DNA. It binds at regions rich in A's and T's (Section 30.10).

Figure 27.9 ▶
The three helical forms of DNA.

AADTYDZ0 missing size

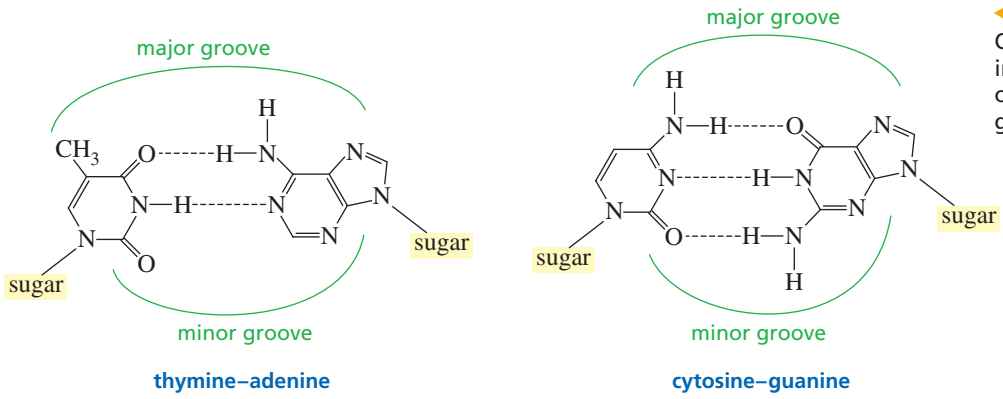


Figure 27.10
One side of each base pair faces into the major groove, and the other side faces into the minor groove.

PROBLEM 13◆

Calculate the length of a turn in:

a. A-DNA b. B-DNA c. Z-DNA

27.9 Biosynthesis of DNA: Replication

3-D Molecules: B-helix; A-helix; Z-helix

Watson and Crick's proposal for the structure of DNA was an exciting development because the structure immediately suggested how DNA is able to pass on genetic information to succeeding generations. Because the two strands are complementary, both carry the same genetic information. Both strands serve as templates for the synthesis of complementary new strands (Figure 27.11). The new (daughter) DNA

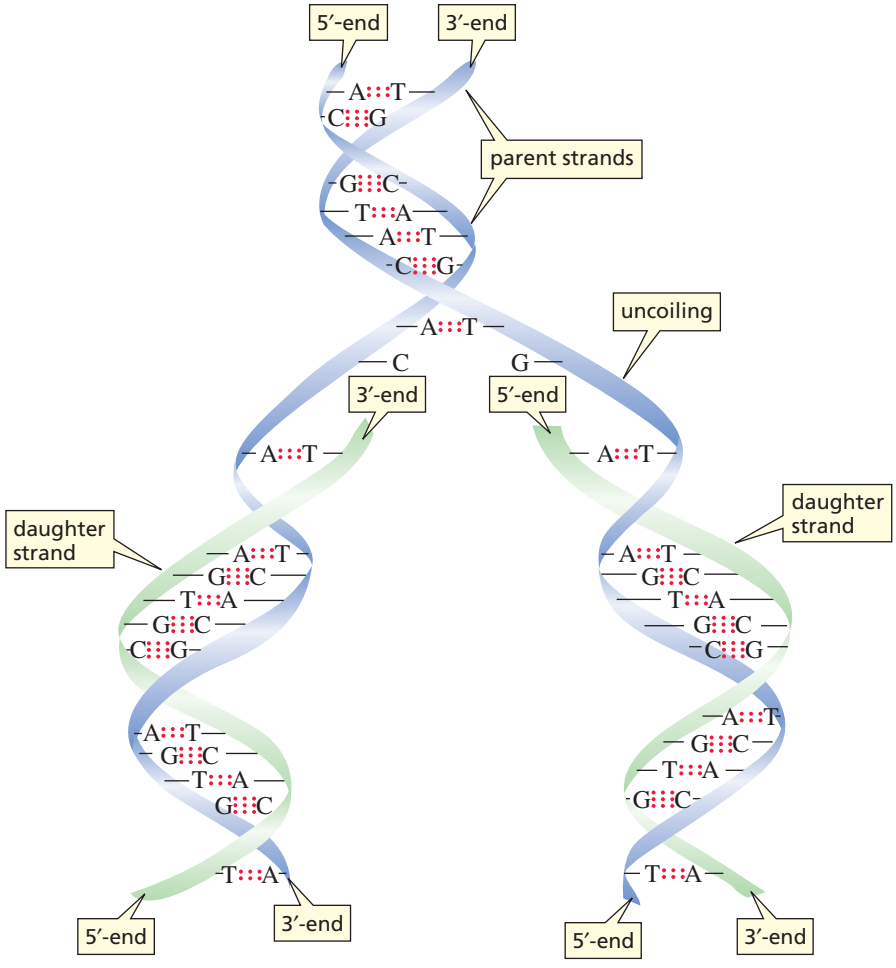


Figure 27.11
Replication of DNA. The daughter strand on the left is synthesized continuously in the 5′ ⟶ 3′ direction; the daughter strand on the right is synthesized discontinuously in the 5′ ⟶ 3′ direction.

molecules are identical to the original (parent) molecule—they contain all the original genetic information. The synthesis of identical copies of DNA is called **replication**.

All reactions involved in nucleic acid synthesis are catalyzed by enzymes. The synthesis of DNA takes place in a region of the molecule where the strands have started to separate, called a **replication fork**. Because a nucleic acid can be synthesized only in the 5′ ⟶ 3′ direction, only the daughter strand on the left in Figure 27.11 is synthesized continuously in a single piece (because it is synthesized in the 5′ ⟶ 3′ direction). The other daughter strand needs to grow in the 3′ ⟶ 5′ direction, so it is synthesized discontinuously in small pieces. Each piece is synthesized in the 5′ ⟶ 3′ direction and the fragments are joined together by an enzyme called DNA ligase. Each of the two resulting daughter molecules of DNA that result contains one of the original strands (blue strand) plus a newly synthesized strand (green strand). This process is called **semiconservative replication**.

The genetic information of a human cell is contained in 23 pairs of chromosomes. Each chromosome is composed of several thousand **genes** (segments of DNA). The total DNA of a human cell—the **human genome**—contains 3.1 billion base pairs.

PROBLEM 14

Using a dark line for parental DNA and wavy lines for DNA synthesized from parental DNA, show what the population of DNA molecules would look like in the fourth generation.

PROBLEM 15◆

Assuming that the human genome, with its 3.1 billion base pairs, is entirely in a B-helix, how long is the DNA in a human cell?

PROBLEM 16

Why doesn't DNA unravel completely before replication begins?

27.10 Biosynthesis of RNA: Transcription

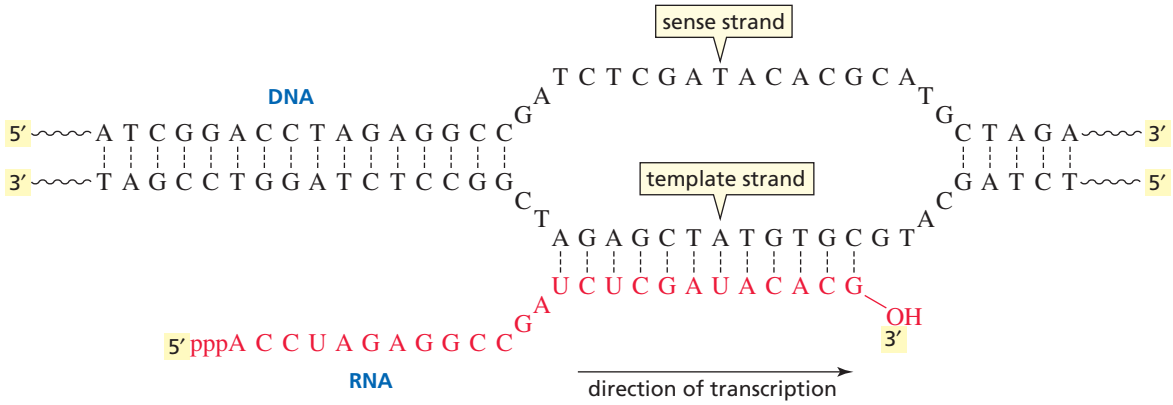
The sequence of DNA bases provides the blueprint for the synthesis of RNA. The synthesis of RNA from a DNA blueprint, called **transcription**, takes place in the nucleus of the cell. This initial RNA is the precursor to all RNA: messenger RNA, ribosomal RNA, and transfer RNA. The newly synthesized RNA leaves the nucleus, carrying the genetic information into the cytoplasm (the cell material outside the nucleus), where translation of this information into proteins takes place (see Figure 27.17).

DNA contains sequences of bases known as *promoter sites*. The **promoter sites** mark the beginning of genes. An enzyme recognizes a promoter site and binds to it, initiating RNA synthesis. The DNA at a promoter site unwinds to give two single strands, exposing the bases. One of the strands is called the **sense strand** or **informational strand**. The complementary strand is called the **template strand** or **antisense strand**. The template strand is read in the 3′ ⟶ 5′ direction, so that RNA can be synthesized in the 5′ ⟶ 3′ direction (Figure 27.12). The bases in the template strand specify the bases that need to be incorporated into RNA, following the same base pairing found in DNA. For example, each guanine in the template strand specifies the incorporation of a cytosine into RNA, and each adenine in the template strand specifies the incorporation of a uracil into RNA. (Recall that in RNA, uracil is used instead of thymine.). Because both the sense strand and RNA are complementary to the template strand, the sense strand and RNA have the same base sequence, except that RNA has a uracil wherever the sense strand has a thymine. Just as there are promoter sites that signal the places to start RNA synthesis, there are sites in DNA that signal that no more bases should be added to the growing strand of RNA, at which point synthesis stops.

Severo Ochoa *was the first to prepare synthetic strands of RNA by incubating nucleotides in the presence of enzymes that are involved in the biosynthesis of RNA.* **Arthur Kornberg** *prepared synthetic strands of DNA in a similar manner. For this work, they shared the 1959 Nobel Prize in physiology or medicine.*

Severo Ochoa (1905–1993) *was born in Spain. He graduated from the University of Malaga in 1921 and received an M.D. from the University of Madrid. He spent the next four years studying in Germany and England and then joined the faculty at New York University College of Medicine. He became a U.S. citizen in 1956.*

Arthur Kornberg *was born in New York in 1918. He graduated from the College of the City of New York and received an M.D. from the University of Rochester. He is a member of the faculty of the biochemistry department at Stanford University.*



▲ **Figure 27.12**
Transcription: using DNA as a blueprint for RNA.

Surprisingly, a gene is not necessarily a continuous sequence of bases. Often the bases of a gene are interrupted by bases that appear to have no informational content. A stretch of bases representing a portion of a gene is called an **exon**, while a stretch of bases that contains no genetic information is called an **intron**. The RNA that is synthesized is complementary to the entire sequence of DNA bases—exons and introns. So after the RNA is synthesized, but before it leaves the nucleus, the so-called nonsense bases (encoded by the introns) are cut out and the informational fragments are spliced together, resulting in a much shorter RNA molecule. This RNA processing step is known as **RNA splicing**. Scientists have found that only about 2% of DNA contains genetic information, while 98% consists of introns.

It has been suggested that the purpose of introns is to make RNA more versatile. The originally synthesized long strand of RNA can be spliced in different ways to create a variety of shorter RNAs.

PROBLEM 17

Why do both thymine and uracil specify the incorporation of adenine?

27.11 Ribosomal RNA

RNA is much shorter than DNA and is generally single-stranded. Although DNA molecules can have billions of base pairs, RNA molecules rarely have more than 10,000 nucleotides. There are three kinds of RNA—**messenger RNA (mRNA)** whose sequence of bases determines the sequence of animo acids in a protein, **ribosomal RNA (rRNA)**, a structural component of ribosomes, and transfer RNA (tRNA), the carriers of amino acids for protein synthesis.

The biosynthesis of proteins takes place on particles known as **ribosomes**. A ribosome is composed of about 40% protein and about 60% rRNA. There is increasing evidence that protein synthesis is catalyzed by rRNA molecules rather than by enzymes. RNA molecules—found in ribosomes—that act as catalysts are known as **ribozymes**. The protein molecules in the ribosome enhance the functioning of the rRNA molecules.

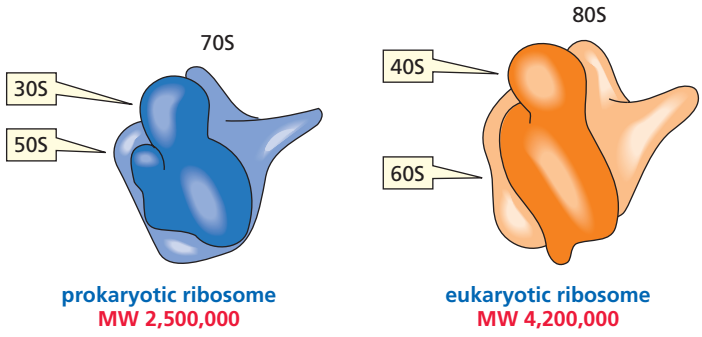
Ribosomes are made up of two subunits. The size of the subunits depends on whether they are found in prokaryotic organisms or eukaryotic organisms. **Prokaryotic organisms** (*pro*, Greek for "before"; *karyon*, Greek for "kernel" or "nut") are the earliest organisms. They are unicellular and do not have nuclei. A **eukaryotic organism** (*eu*, Greek for "well") is much more complicated. Eukaryotic organisms can be unicellular or multicellular and their cells have nuclei. A prokaryotic ribosome is composed of a 50S subunit and a smaller 30S subunit; together they form a 70S ribosome. A eukaryotic ribosome has a 60S subunit and a 40S subunit; together they form an 80S ribosome.

Sidney Altman *and* **Thomas R. Cech** *received the 1989 Nobel Prize in chemistry for their discovery of the catalytic properties of RNA.*

Sidney Altman *was born in Montreal in 1939. He received a B.S. from MIT and a Ph.D. from the University of Colorado, Boulder. He was a postdoctoral fellow in Francis Crick's laboratory at Cambridge University. He is a professor of biology at Yale University.*

Thomas Cech *was born in Chicago in 1947. He received a B.A. from Grinnell College and a Ph.D. from the University of California, Berkeley. He was a postdoctoral fellow at MIT. He is a professor of chemistry at the University of Colorado, Boulder.*

The S stands for the **sedimentation constant**, which designates where a given component sediments during centrifugation.[3]



27.12 Transfer RNA

Transfer RNA (tRNA) is much smaller than mRNA or rRNA. It contains only 70 to 90 nucleotides. The single strand of tRNA is folded into a characteristic cloverleaf structure strung out with three loops and a little bulge next to the right-hand loop (Figure 27.13a). There are at least four regions with complementary base pairing. All tRNAs have a CCA sequence at the 3′-end. The three bases at the bottom of the loop directly opposite the 5′- and 3′-ends are called an **anticodon** (Figures 27.13a and b).

Each tRNA can carry an amino acid bound as an ester to its terminal 3′-OH group. The amino acid will be inserted into a protein during protein biosynthesis. Each tRNA can carry only one particular amino acid. A tRNA that carries alanine is designated as $tRNA^{Ala}$.

3-D Molecule: tRNA

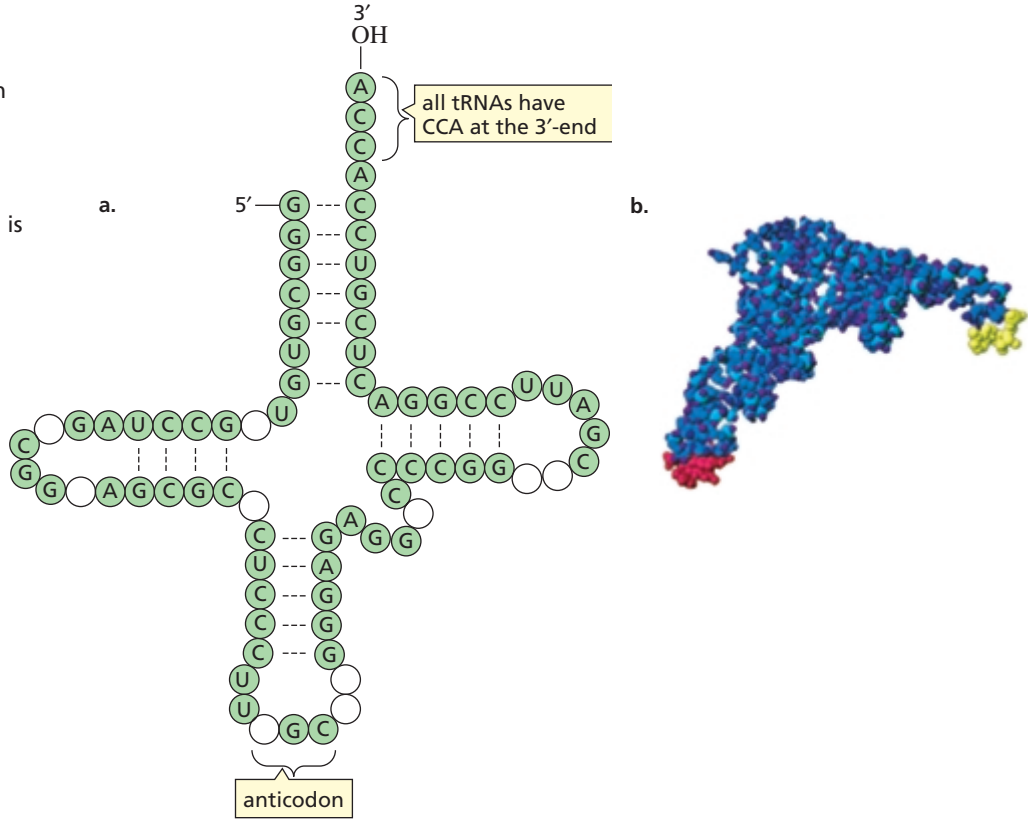
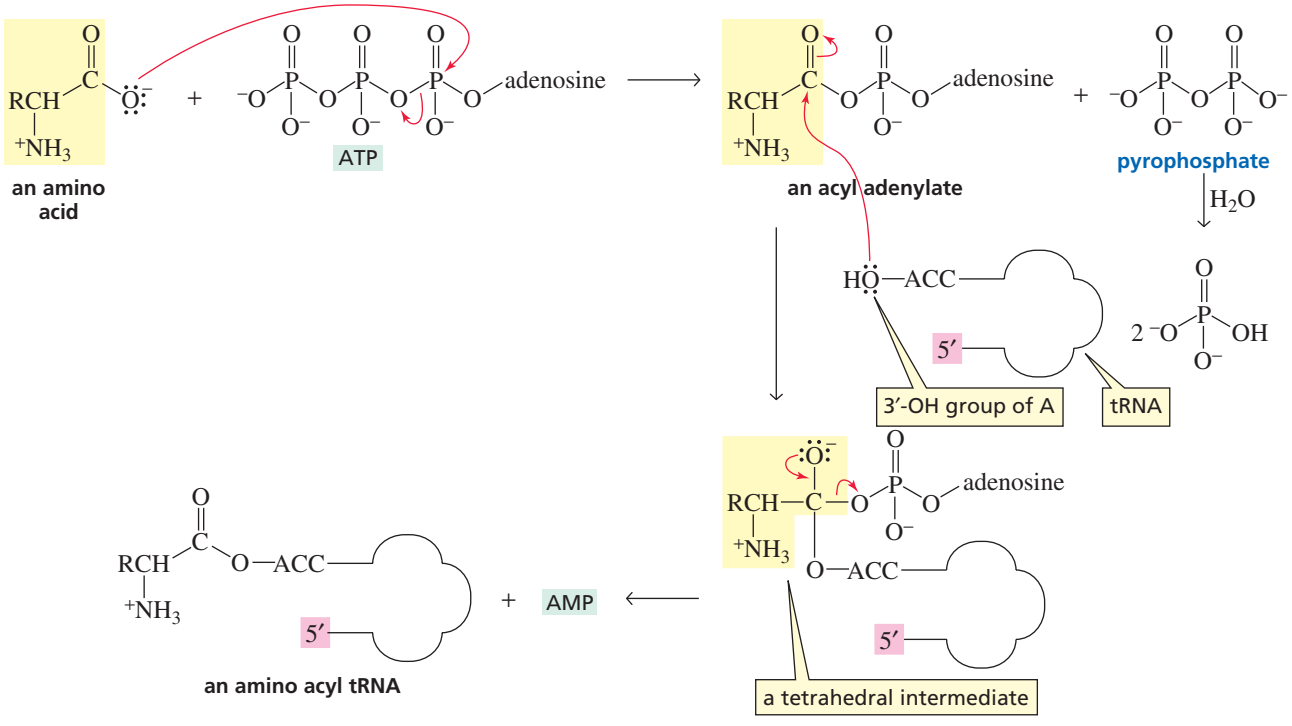


Figure 27.13 ▶
(a) $tRNA^{Ala}$, a transfer RNA that carries alanine. Compared with other RNAs, tRNA contains a high percentage of unusual bases (shown as empty circles). These bases result from enzymatic modification of the four normal bases. (b) $tRNA^{Ser}$: The anticodon is shown in red; the serine binding site is shown in yellow.

[3]Sedimentation constants are not additive, which is why a 50S and a 30S can combine to form a 70S.

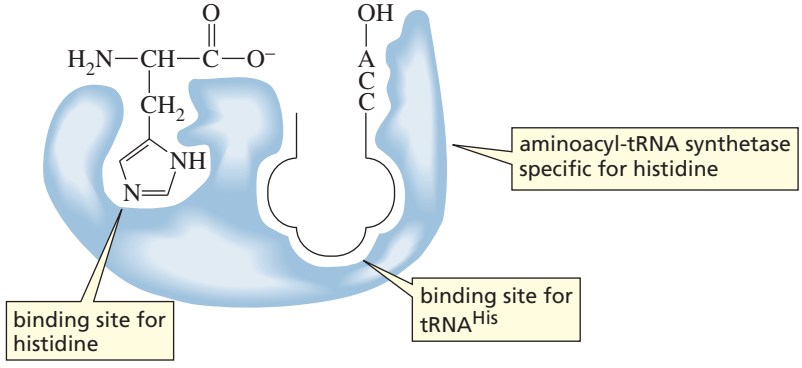
How does an amino acid become attached to a tRNA? Attachment of the amino acid is catalyzed by an enzyme called aminoacyl-tRNA synthetase. In the first step of the enzyme-catalyzed reaction (Figure 27.14), the carboxyl group of the amino acid attacks the α-phosphorus of ATP, activating the carboxyl group by forming an acyl adenylate. The pyrophosphate that is expelled is subsequently hydrolyzed, ensuring the irreversibility of the phosphoryl transfer reaction (Section 27.3). Then a nucleophilic acyl substitution reaction occurs—the 3′-OH group of tRNA attacks the carbonyl carbon of the acyl adenylate, forming a tetrahedral intermediate. The amino acyl tRNA is formed when AMP is expelled from the tetrahedral intermediate. All the steps take place at the active site of the enzyme. Each amino acid has its own aminoacyl-tRNA synthetase. Each synthetase has two specific binding sites, one for the amino acid and one for the tRNA that carries that amino acid (Figure 27.15).

Elizabeth Keller (1918–1997) *was the first to recognize that tRNA had a cloverleaf structure. She received a B.S. from the University of Chicago in 1940 and a Ph.D. from Cornell University Medical College in 1948. She worked at the Huntington Memorial Laboratory of Massachusetts General Hospital and at the United States Public Health Service. Later she became a professor at MIT and then at Cornell University.*



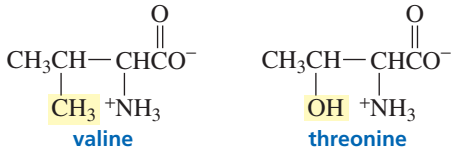
▲ **Figure 27.14**
The proposed mechanism for aminoacyl-tRNA synthetase—the enzyme that catalyzes the attachment of an amino acid to a tRNA.

It is critical that the correct amino acid is attached to the tRNA. Otherwise, the correct protein will not be synthesized. Fortunately, the synthetases correct their own mistakes. For example, valine and threonine are approximately the same size—threonine has an OH group in place of a CH_3 group of valine. Both amino acids, therefore, can bind at the amino acid binding site of the aminoacyl-tRNA synthetase for valine, and both can then be activated by reacting with ATP to form an acyl adenylate. The aminoacyl-tRNA synthetase for valine has two adjacent catalytic sites, one for



◀ **Figure 27.15**
An aminoacyl-tRNA synthetase has a binding site for tRNA and a binding site for the particular amino acid that is to be attached to that tRNA. Histidine is the amino acid and $tRNA^{His}$ is the tRNA molecule in this example.

attaching the acyl adenylate to tRNA and one for hydrolyzing the acyl adenylate. The acylation site is hydrophobic, so valine is preferred over threonine for the tRNA acylation reaction. The hydrolytic site is polar, so threonine is preferred over valine for the hydrolysis reaction. Thus, if threonine is activated by the aminoacyl-tRNA synthetase for valine, it will be hydrolyzed rather than transferred to the tRNA.



27.13 Biosynthesis of Proteins: Translation

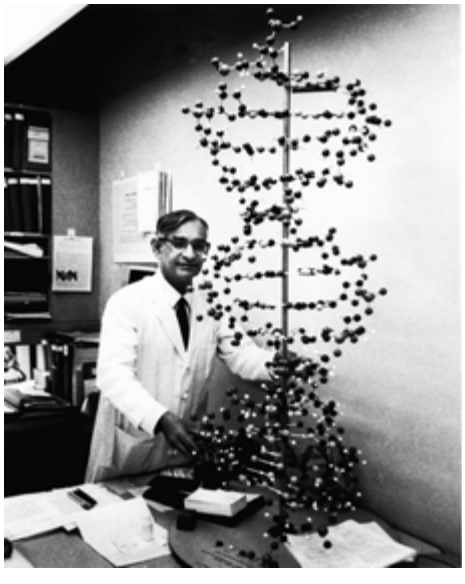
A protein is synthesized from its N-terminal end to its C-terminal end by reading the bases along the mRNA strand in the 5′ ⟶ 3′ direction. A sequence of three bases, called a **codon**, specifies a particular amino acid that is to be incorporated into a protein. The bases are read consecutively and are never skipped. A codon is written with the 5′-nucleotide on the left. Each amino acid is specified by a three-base sequence known as the **genetic code** (Table 27.2). For example, UCA on mRNA codes for the amino acid serine, whereas CAG codes for glutamine.

Because there are four bases and the codons are triplets, $4^3 = 64$ different codons are possible. This is many more than are needed to specify the 20 different amino acids, so all the amino acids—except methionine and tryptophan—have more than one codon. It is not surprising, therefore, that methionine and tryptophan are the least abundant amino acids in proteins. Actually, 61 of the bases specify amino acids, and three bases are stop codons. **Stop codons** tell the cell to "stop protein synthesis here."

Translation is the process by which the genetic message in DNA that has been passed to mRNA is decoded and used to build proteins. Each of the approximately 100,000 proteins in the human body is synthesized from a different mRNA. Don't

The genetic code was worked out independently by **Marshall Nirenberg** *and* **Har Gobind Khorana**, *for which they shared the 1968 Nobel Prize in physiology or medicine.* **Robert Holley**, *who worked on the structure of tRNA molecules, also shared that year's prize.*

Marshall Nirenberg *was born in New York in 1927. He received a bachelor's degree from the University of Florida and a Ph.D. from the University of Michigan. He is a scientist at the National Institutes of Health.*



Har Gobind Khorana *was born in India in 1922. He received a bachelor's and a master's degree from Punjab University and a Ph.D. from the University of Liverpool. In 1960 he joined the faculty at the University of Wisconsin and later became a professor at MIT.*

Table 27.2 The Genetic Code

5′-Position	Middle position				5′-Position
	U	C	A	G	
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	Stop	Stop	A
	Leu	Ser	Stop	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

confuse transcription and translation—these words are used just as they are used in English. Transcription (DNA to RNA) is copying *within the same language* of nucleotides. Translation (RNA to protein) is *changing to another language*—the language of amino acids.

Transcription: DNA ⟶ RNA

Translation: mRNA ⟶ protein

How the information in mRNA is translated into a polypeptide is shown in Figure 27.16. In this figure, serine was the last amino acid incorporated into the

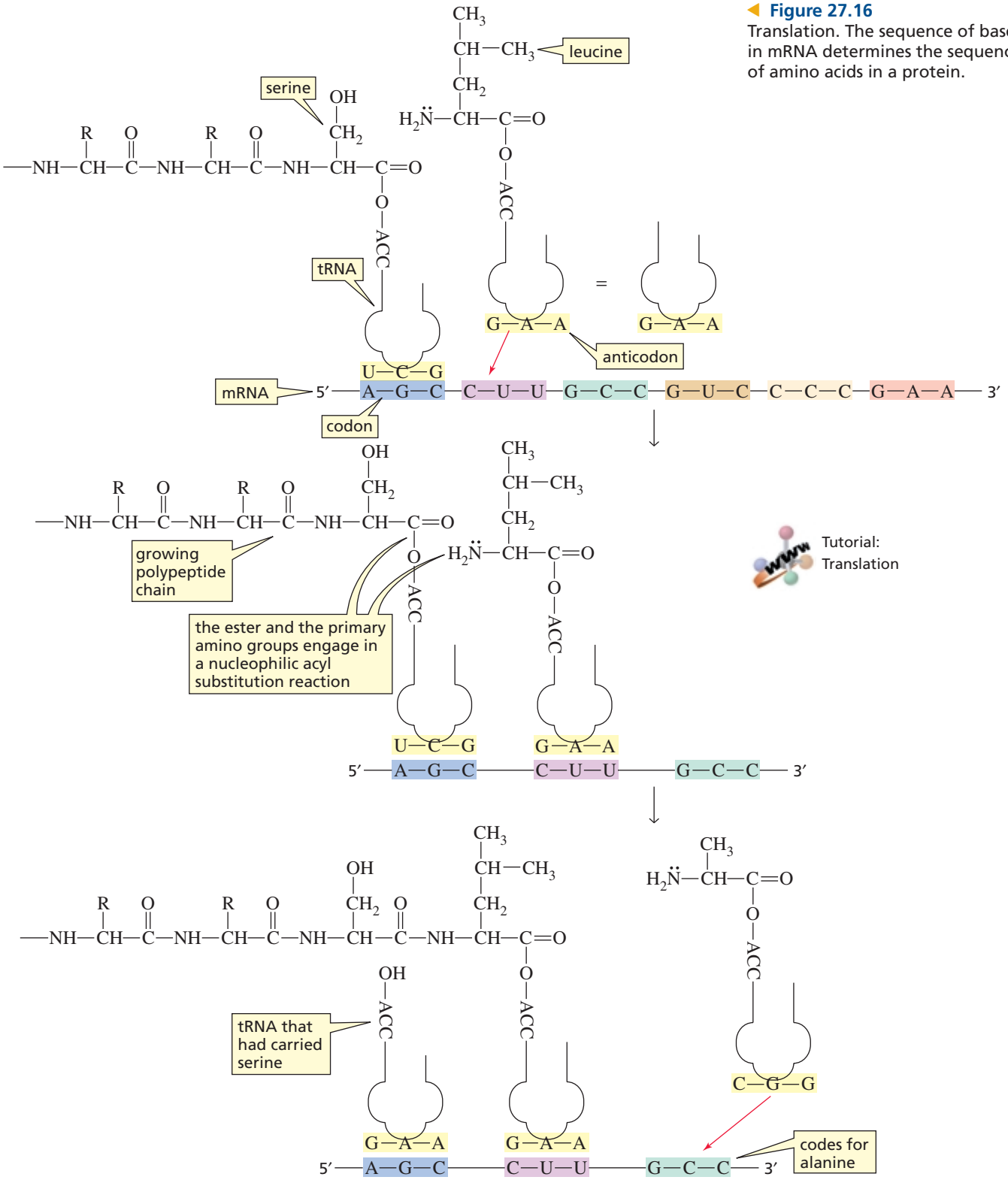


Figure 27.16
Translation. The sequence of bases in mRNA determines the sequence of amino acids in a protein.

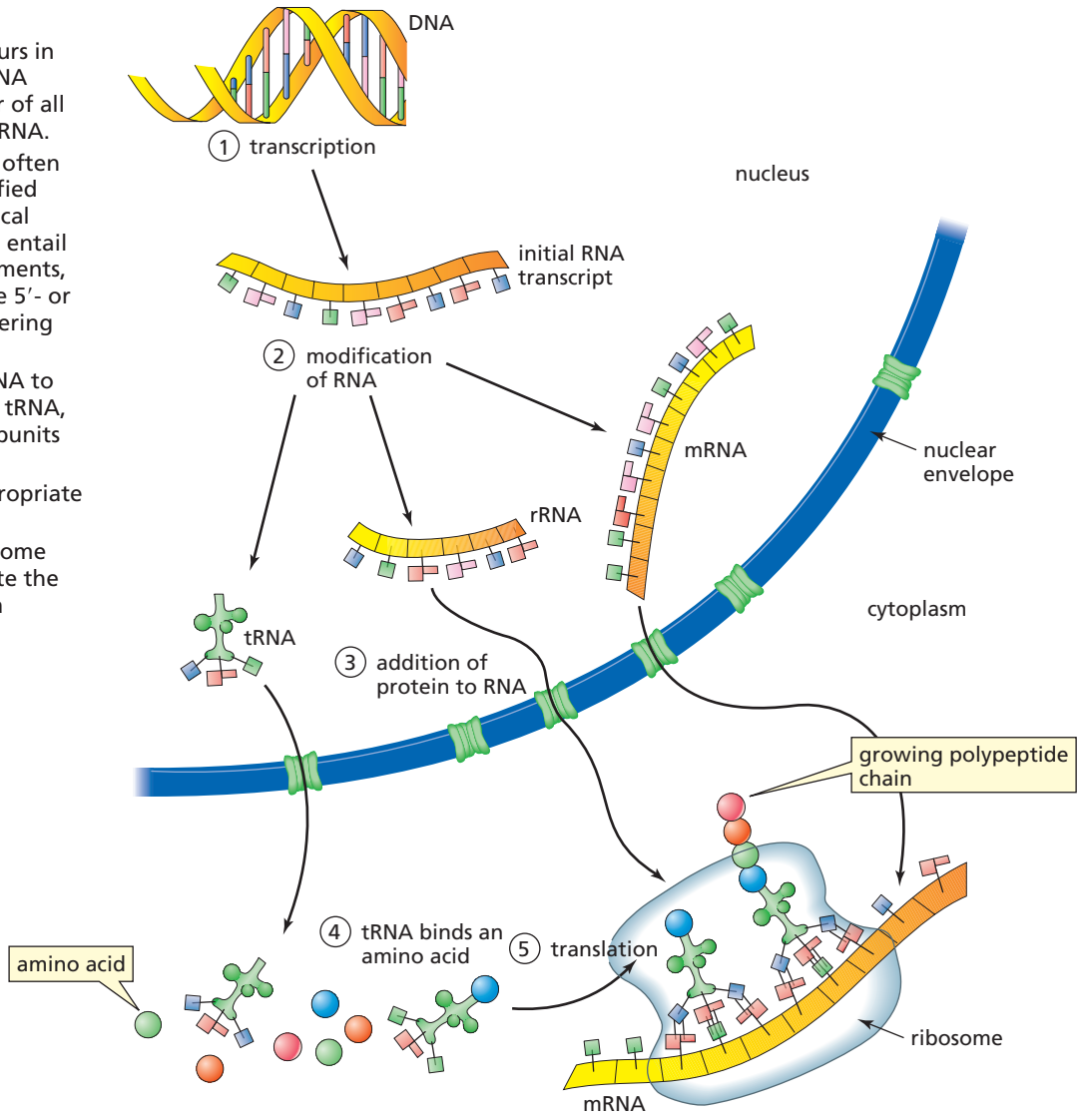
Tutorial: Translation

growing polypeptide chain. Serine was specified by the AGC codon because the anticodon of the tRNA that carries serine is GCU (3′-UCG-5′). (Remember that a base sequence is read in the 5′ ⟶ 3′ direction, so the sequence of bases in an anticodon must be read from right to left.) The next codon is CUU, signaling for a tRNA with an anticodon of AAG (3′-GAA-5′). That particular tRNA carries leucine. The amino group of leucine reacts in an enzyme-catalyzed nucleophilic acyl substitution reaction with the ester on the adjacent tRNA, displacing the tRNA. The next codon (GCC) brings in a tRNA carrying alanine. The amino group of alanine displaces the tRNA that brought in leucine. Subsequent amino acids are brought in one at a time in the same way, with the codon in mRNA specifying the amino acid to be incorporated by complementary base pairing with the anticodon of the tRNA that carries that amino acid.

Protein synthesis takes place on the ribosomes (Figure 27.17). The smaller subunit of the ribosome (30S in prokaryotic cells) has three binding sites for RNA molecules. It binds the mRNA whose base sequence is to be read, the tRNA carrying the growing peptide chain, and the tRNA carrying the next amino acid to be incorporated into the protein. The larger subunit of the ribosome (50S in prokaryotic cells) catalyzes peptide bond formation.

Figure 27.17 ▶

1. Transcription of DNA occurs in the nucleus. The initial RNA transcript is the precursor of all RNA: tRNA, rRNA, and mRNA.
2. The initially formed RNA often must be chemically modified before it acquires biological activity. Modification can entail removing nucleotide segments, adding nucleotides to the 5′- or 3′-ends, or chemically altering certain nucleotides.
3. Proteins are added to rRNA to form ribosomal subunits. tRNA, mRNA, and ribosomal subunits leave the nucleus.
4. Each tRNA binds the appropriate amino acid.
5. tRNA, mRNA, and a ribosome work together to translate the mRNA information into a protein.



PROBLEM 18◆

If methionine is the first amino acid incorporated into a heptapeptide, what is the sequence of the amino acids encoded for by the following stretch of mRNA?

5′—G—C—A—U—G—G—A—C—C—C—C—G—U—U—A—U—U—A—A—A—C—A—C—3′

PROBLEM 19◆

Four C's occur in a row in the segment of mRNA in Problem 18. What polypeptide would be formed from the mRNA if one of the four C's were cut out of the strand?

PROBLEM 20

UAA is a stop codon. Why does the UAA sequence in mRNA in Problem 18 not cause protein synthesis to stop?

PROBLEM 21◆

Write the sequences of bases in the sense strand of DNA that resulted in the mRNA in Problem 18.

PROBLEM 22

List the possible codons on mRNA that specify each amino acid in Problem 18 and the anticodon on the tRNA that carries that amino acid.

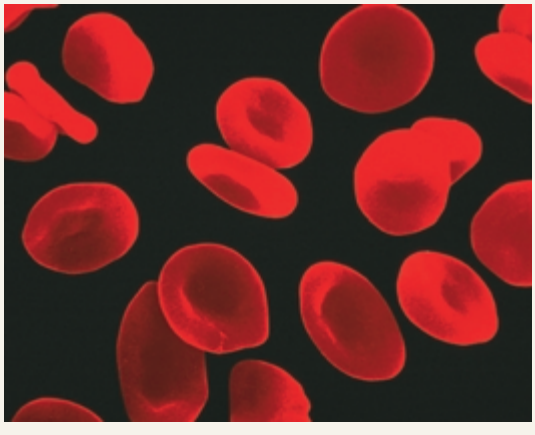
3-D Molecules: Chloramphenicol complexed to acetyl-transferase; Tetracycline

A sculpture done by Robert Holley.

Robert W. Holley (1922–1993) *was born in Illinois and received a bachelor's degree from the University of Illinois and a Ph.D. from Cornell University. During World War II he worked on the synthesis of penicillin at Cornell Medical School. He was a professor at Cornell and later at the University of California, San Diego. He was also a noted sculptor.*

SICKLE CELL ANEMIA

Sickle cell anemia is an example of the damage that can be caused by a change in a single base of DNA (Problem 55 in Chapter 23). It is a hereditary disease caused when a GAG triplet becomes a GTG triplet in the sense strand of a section of DNA that codes for the β-subunit of hemoglobin. As a consequence, the mRNA codon becomes GUG—which signals for incorporation of valine—rather than GAG, which would have signaled for incorporation of glutamic acid. The change from a polar glutamic acid to a nonpolar valine is sufficient to change the shape of the deoxyhemoglobin molecule and induce aggregation, causing it to precipitate in red blood cells. This stiffens the cells, making it difficult for them to squeeze through a capillary. Blocked capillaries cause severe pain and can be fatal.

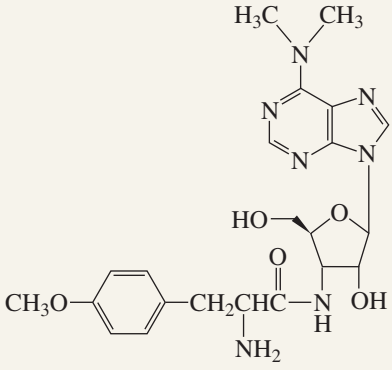


Normal red blood cells

Sickle red blood cells

ANTIBIOTICS THAT ACT BY INHIBITING TRANSLATION

Puromycin is a naturally occurring antibiotic. It is one of several antibiotics that act by inhibiting translation. Puromycin mimics the 3′-CCA-aminoacyl portion of a tRNA. If, during translation, the enzyme is fooled into transferring the growing peptide chain to the amino group of puromycin rather than to the amino group of the incoming 3′-CCA-aminoacyl tRNA, protein synthesis stops. Because puromycin blocks protein synthesis in eukaryotes as well as in prokaryotes, it is poisonous to humans and therefore is not a clinically useful antibiotic. To be clinically useful, an antibiotic must affect protein synthesis only in prokaryotic cells.

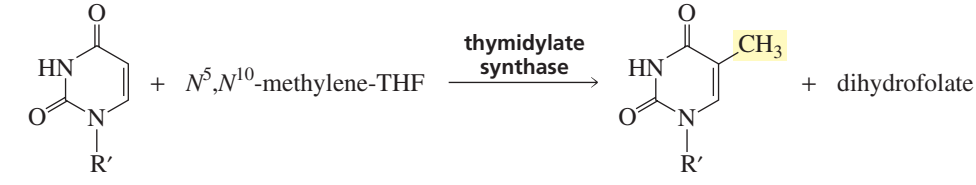


puromycin

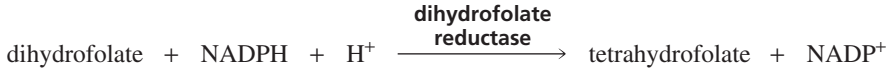
Clinically useful antibiotics	Mode of action
Tetracycline	Prevents the aminoacyl-tRNA from binding to the ribosome
Erythromycin	Prevents the incorporation of new amino acids into the protein
Streptomycin	Inhibits the initiation of protein synthesis
Chloramphenicol	Prevents the new peptide bond from being formed

27.14 Why DNA Contains Thymine Instead of Uracil

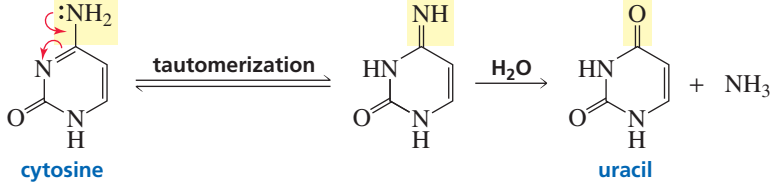
In Section 25.8 we saw that dTMP is formed when dUMP is methylated, with coenzyme N^5,N^{10}-methylenetetrahydrofolate supplying the methyl group. Because the incorporation of the methyl group into uracil oxidizes tetrahydrofolate to dihydrofolate, dihydrofolate must be reduced back to tetrahydrofolate to prepare the cofactor for another catalytic cycle. The reducing agent is NADPH. Every NADPH formed in a biological organism can drive the formation of three ATPs, so using an NADPH to reduce dihydrofolate comes at the expense of ATP. This means that the synthesis of thymine is energetically expensive, so there must be a good reason for DNA to contain thymine instead of uracil.



dUMP
R′ = 2′-deoxyribose-5-P
dTMP



The presence of thymine instead of uracil in DNA prevents potentially lethal mutations. Cytosine can tautomerize to form an imine, which can be hydrolyzed to uracil (Section 18.6). The overall reaction is called a **deamination** since it removes an amino group.



cytosine
uracil

If a cytosine in DNA is deaminated to a uracil, uracil will specify incorporation of an adenine into the daughter strand during replication instead of the guanine that would have been specified by cytosine. Fortunately, a U in DNA is recognized as a "mistake" by cell enzymes before an incorrect base can be inserted into the daughter strand. These enzymes cut out the U and replace it with a C. If U's were normally found in DNA, the enzymes could not distinguish between a normal U and a U formed by deamination of cytosine. Having T's in place of U's in DNA allows the U's that are found in DNA to be recognized as mistakes.

Unlike DNA, which replicates itself, any mistake in RNA does not survive for long because RNA is constantly being degraded and then resynthesized from the DNA template. Therefore, it is not worth spending the extra energy to incorporate T's into RNA.

PROBLEM 23◆

Adenine can be deaminated to hypoxanthine, and guanine can be deaminated to xanthine. Draw structures for hypoxanthine and xanthine.

PROBLEM 24

Explain why thymine cannot be deaminated.

27.15 Determining the Base Sequence of DNA

In June 2000, two teams of scientists (one from a private biotechnology company and one from the publicly funded Human Genome Project) announced that they had completed the first draft of the sequence of the 3.1 billion base pairs in human DNA. This is an enormous accomplishment. For example, if the sequence of 1 million base pairs were determined each day, it would take more than 10 years to complete the sequence of the human genome.

DNA molecules are too large to sequence as a unit, so DNA is first cleaved at specific base sequences and the resulting DNA fragments are sequenced. The enzymes that cleave DNA at specific base sequences are called **restriction endonucleases**, and the DNA fragments that are formed are called **restriction fragments**. Several hundred restriction enzymes are now known. A few examples of restriction enzymes, the base sequence each recognizes, and the point of cleavage in that base sequence are shown here.

restriction enzyme	recognition sequence
*Alu*I	AG↓CT TC↑GA
*Fnu*DI	GG↓CC CC↑GG
*Pst*I	CTGCA↓G G↑ACGTC

The base sequences that most restriction enzymes recognize are *palindromes*. A palindrome is a word or a group of words that reads the same forward and backward. "Toot" and "race car" are examples of palindromes.[4] A restriction enzyme recognizes

[4]Some other palindromes are "Mom," "Dad," "Bob," "Lil," "radar," "noon," "wow," "poor Dan in a droop", "a man, a plan, a canal, Panama," "Sex at noon taxes," and "He lived as a devil, eh?"

a piece of DNA in which *the template strand is a palindrome of the sense strand*. In other words, the sequence of bases in the template strand (reading from right to left) is identical to the sequence of bases in the sense strand (reading from left to right).

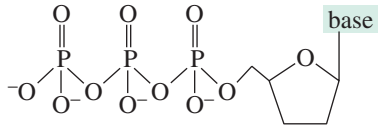
PROBLEM 25◆

Which of the following base sequences would most likely be recognized by a restriction endonuclease?

a. ACGCGT
b. ACGGGT
c. ACGGCA
d. ACACGT
e. ACATCGT
f. CCAACC

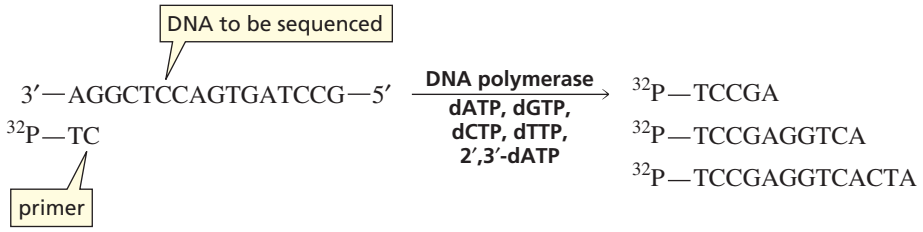
The restriction fragments can be sequenced using a chain-terminator procedure developed by Frederick Sanger known as the **dideoxy method**. This method involves generating fragments whose length depends on the last base added to the fragment. Because of its simplicity, it has superceded alternative methods.

In the dideoxy method a small piece of DNA called a primer, labeled at the 5′-end with ^{32}P, is added to the restriction fragment whose sequence is to be determined. Next, the four 2′-deoxyribonucleoside triphosphates are added as well as DNA polymerase, the enzyme that adds nucleotides to a strand of DNA. In addition, a small amount of the 2′,3′-dideoxynucleoside triphosphate of one of the bases is added to the reaction mixture. A 2′,3′-dideoxynucleoside triphosphate has no OH groups at the 2′- and 3′-positions.



a 2′,3′-dideoxynucleoside triphosphate

Nucleotides will be added to the primer by base pairing with the restriction fragment. Synthesis will stop if the 2′,3′-dideoxy analog of dATP is added instead of dATP, because the 2′,3′-dideoxy analog does not have a 3′-OH to which additional nucleotides can be added. Therefore, three different chain-terminated fragments will be obtained from the DNA restriction fragment shown here.



The procedure is repeated three more times using a 2′,3′-dideoxy analog of dGTP, then a 2′,3′-dideoxy analog of dCTP, and then a 2′,3′-dideoxy analog of dTTP.

PROBLEM 26◆

What labeled fragments would be obtained from the segment of DNA shown above if a 2′,3′-dideoxy analog of dGTP had been added to the reaction mixture instead of a 2′,3′-dideoxy analog of dATP?

Frederick Sanger *(Section 23.12) and* **Walter Gilbert** *shared half of the 1980 Nobel Prize in chemistry for their work on DNA sequencing. The other half went to* **Paul Berg**, *who had developed a method of cutting nucleic acids at specific sites and recombining the fragments in new ways, a technique known as recombinant DNA technology.*

Walter Gilbert *was born in Boston in 1932. He received a master's degree in physics from Harvard and a Ph.D. in mathematics from Cambridge University. In 1958 he joined the faculty at Harvard, where he became interested in molecular biology.*

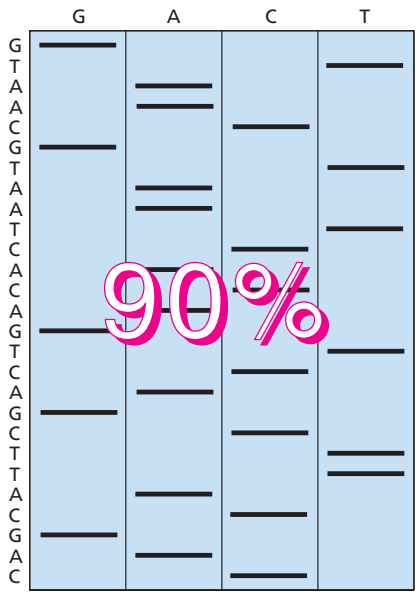
Paul Berg *was born in New York in 1926. He received a Ph.D. from Western Reserve University (now Case Western Reserve University). He joined the faculty at Washington University in St. Louis in 1955 and became a professor of biochemistry at Stanford in 1959.*

The chain-terminated fragments obtained from each of the four experiments are loaded onto separate lanes of a buffered polyacrylamide gel—the fragments obtained from using a 2′,3′-dideoxy analog of dATP are loaded onto one lane, the fragments obtained from using a 2′,3′-dideoxy analog of dGTP onto another lane, and so on. An electric field is applied across the ends of the gel, causing the negatively charged fragments to travel toward the positively charged electrode (the anode). The smaller fragments fit through the spaces in the gel relatively easily and therefore travel through the gel faster, while the larger fragments pass through the gel more slowly.

After the fragments have been separated, the gel is placed in contact with a photographic plate. Radiation from ^{32}P causes a dark spot to appear on the plate opposite the location of each labeled fragment in the gel. This technique is called autoradiography, and the exposed photographic plate is known as an **autoradiograph** (Figure 27.18).

The sequence of bases in the original restriction fragment can be read directly from the autoradiograph. The identity of each base is determined by noting the column where each successive dark spot (larger piece of labeled fragment) appears, starting at the bottom of the gel. The sequence of the fragment of DNA responsible for the autoradiograph in Figure 27.18 is shown on the left-hand side of the figure.

Once the sequence of bases in a restriction fragment is determined, the results can be checked by determining the base sequence of the complementary strand. The base sequence in the original piece of DNA can be determined by repeating the entire procedure with a different restriction endonuclease and noting overlapping fragments.



▲ **Figure 27.18**
An autoradiograph.

DNA FINGERPRINTING

The base sequence of the human genome varies from individual to individual, generally by a single base change every few hundred base pairs. Because some of these changes occur in base sequences recognized by restriction endonucleases, the fragments formed when human DNA reacts with a particular restriction endonuclease vary in size depending on the individual. It is this variation that forms the basis of DNA fingerprinting (also called DNA profiling or DNA typing). This technique is used by forensic chemists to compare DNA samples collected at the scene of a crime with the DNA of the suspected perpetrator. The most powerful technique for DNA identification analyzes restriction fragment length polymorphisms (RFLPs) obtained from regions of DNA in which individual variations are most common. This technique takes four to six weeks and requires a blood stain about the size of a dime. The chance of identical results from two different persons is thought to be one in a million. The second type of DNA profiling uses a polymerase chain reaction (PCR), which amplifies a specific region of DNA and compares differences at that site among individuals. This technique can be done in less than a week and requires only 1% of the amount required for RFLP, but does not discriminate as well among individuals. The chance of identical results from two different people is 1 in 500 to 1 in 2000. DNA fingerprinting is also being used to establish paternity, accounting for about 100,000 DNA profiles a year.

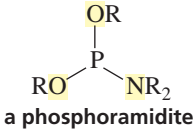
27.16 Laboratory Synthesis of DNA Strands

There is a great deal of interest in the synthesis of oligonucleotides with specific base sequences. This would allow scientists to synthesize genes that could be inserted into the DNA of microorganisms, causing the organisms to synthesize a particular protein. Alternatively, a synthetic gene could be inserted into the DNA of an organism defective in that gene—a process known as **gene therapy**.

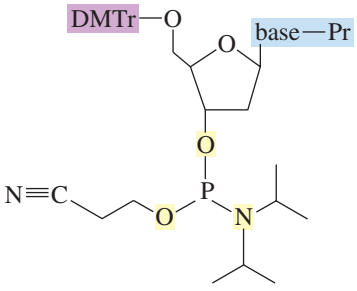
Synthesizing an oligonucleotide with a particular base sequence is an even more challenging task than synthesizing a polypeptide with a specific amino acid sequence because each nucleotide has several groups that must be protected and then deprotected at the proper times. The approach taken was to develop an automated method similar to automated peptide synthesis (Section 23.10). The growing nucleotide is attached to a solid support so that it can be purified by flushing the reaction container with an appropriate solvent. Therefore, none of the synthesized product will be lost during purification.

Phosphoramidite Monomers

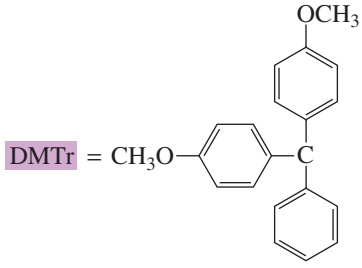
One current method synthesizes oligonucleotides using phosphoramidite monomers. The 5′-OH group of each phosphoramidite monomer is attached to a *para*-dimethoxytrityl (DMTr) protecting group. The particular group (Pr) used to protect the base depends on the base.



a phosphoramidite

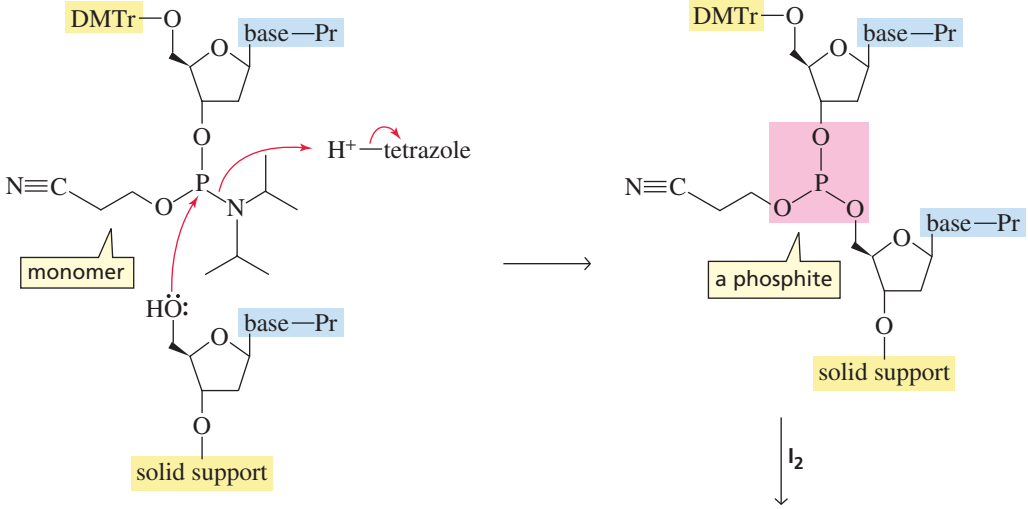


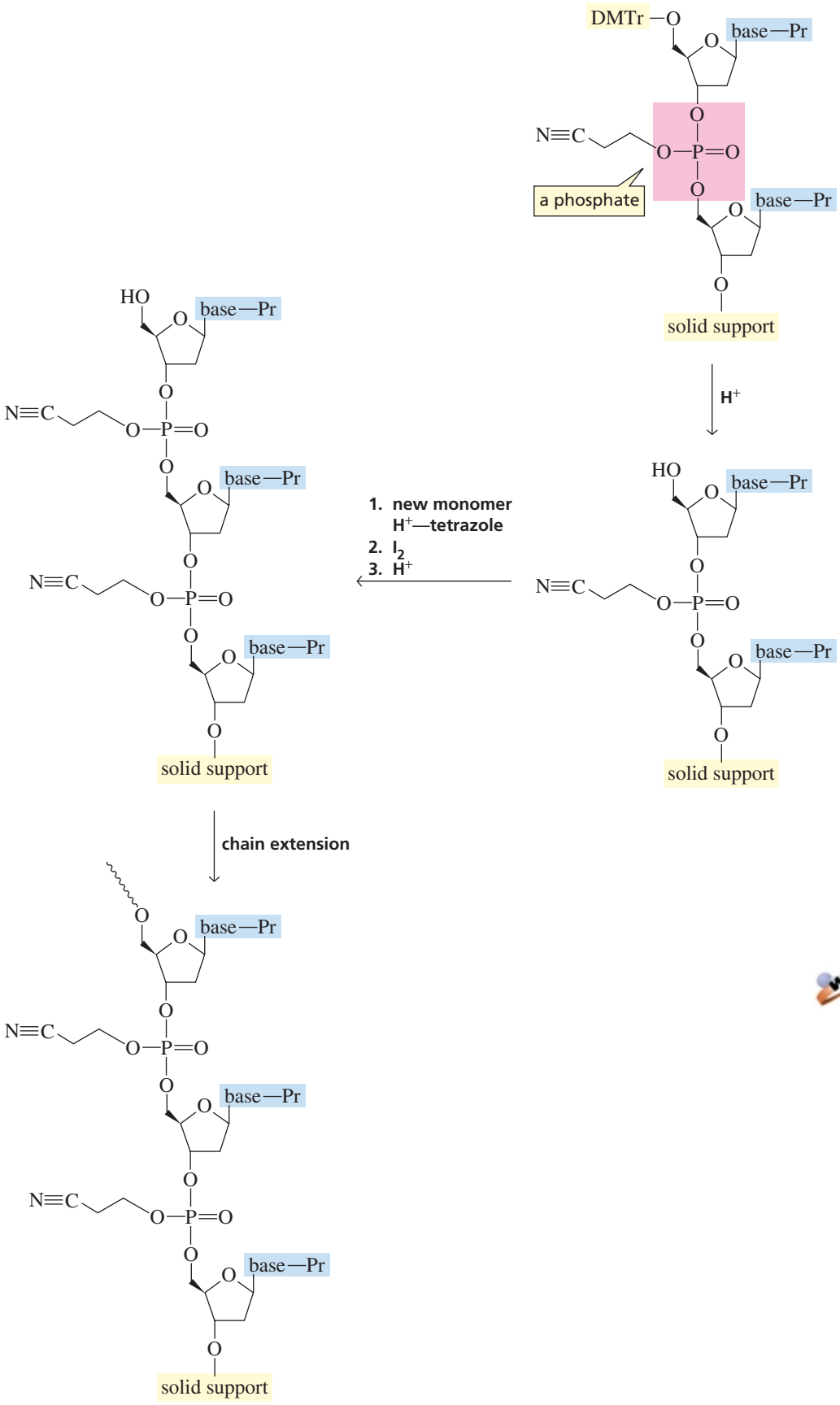
a phosphoramidite monomer used for oligonucleotide synthesis




3-D Molecule: Phosphoramidite

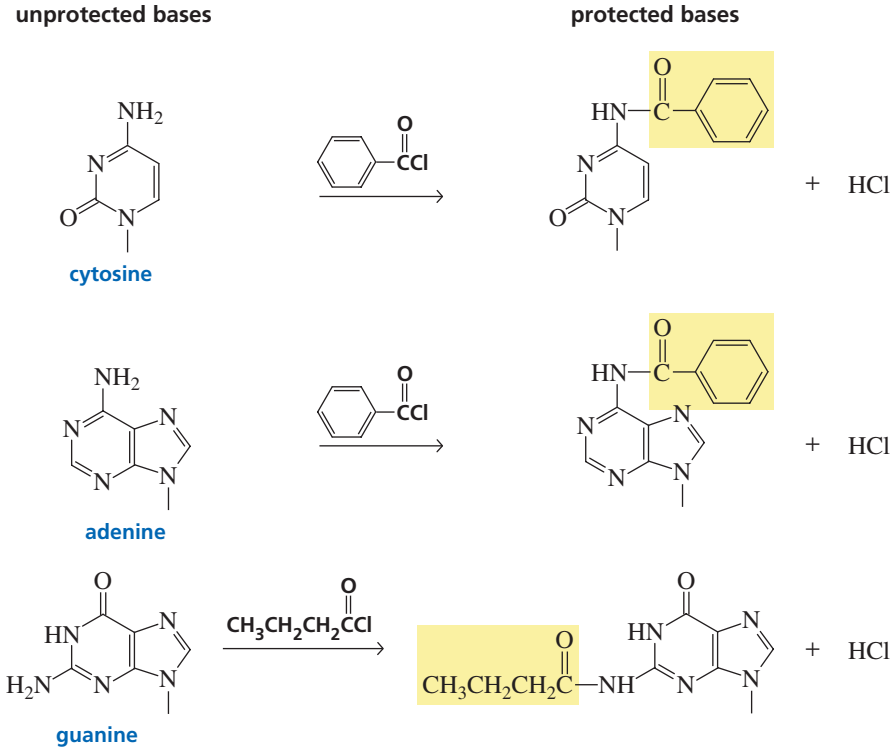
The 3′-nucleoside of the oligonucleotide to be synthesized is attached to a controlled-pore glass solid support and the oligonucleotide is synthesized from the 3′-end. When a monomer is added to the nucleoside attached to the solid support, the only nucleophile in the reaction mixture is the 5′-OH group of the sugar bonded to the solid support. This nucleophile attacks the phosphorus of the phosphoramidite, displacing the amine and forming a phosphite. The amine is too strong a base to be expelled without being protonated. Protonated tetrazole is the acid used for protonation because it is strong enough to protonate the diisopropylamine leaving group, but not strong enough to remove the DMTr protecting group. The phosphite is oxidized to a phosphate using I_2 or *tert*-butylhydroperoxide. The DMTr protecting group on the 5′-end of the dinucleotide is removed with mild acid. The cycle of (1) monomer addition, (2) oxidation, and (3) deprotection with acid is repeated over and over until a polymer of the desired length and sequence is obtained. Notice that the DNA polymer is synthesized in the 3′ ⟶ 5′ direction, which is opposite to the direction (5′ ⟶ 3′) DNA is synthesized in nature.



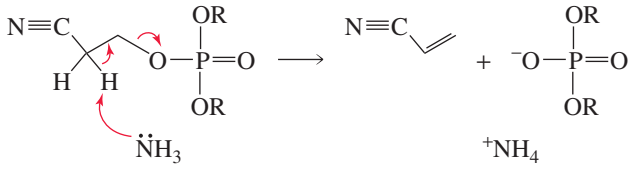


Tutorial:
Oligonucleotide synthesis
with phosphoramidites

The NH_2 groups of cytosine, adenine, and guanine are nucleophiles and therefore must be protected to prevent them from reacting with a newly added monomer. The NH_2 groups of cytosine and adenine are protected as benzamides and the NH_2 group of guanine is protected as a butanamide. Thymine does not contain any nucleophilic groups, so it does not have to be protected.



After the oligonucleotide is synthesized, the protecting groups on the phosphates and the protecting groups on the bases must be removed, and the oligonucleotide has to be detached from the solid support. This can all be done in a single step using aqueous ammonia. Because a hydrogen bonded to a carbon adjacent to a cyano group is relatively acidic (Section 19.1), ammonia can be used as the base in the elimination reaction that removes the phosphate protecting group.

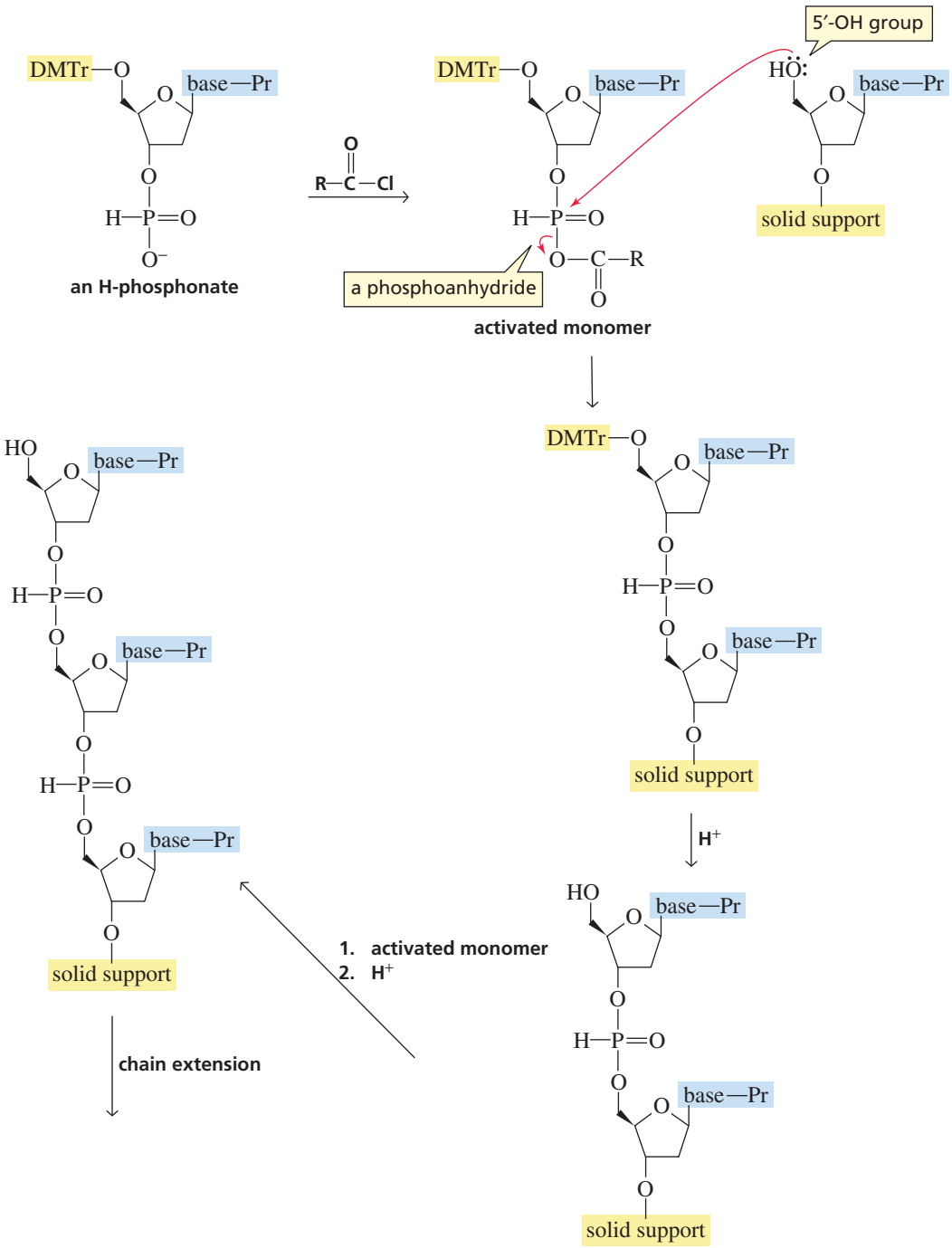


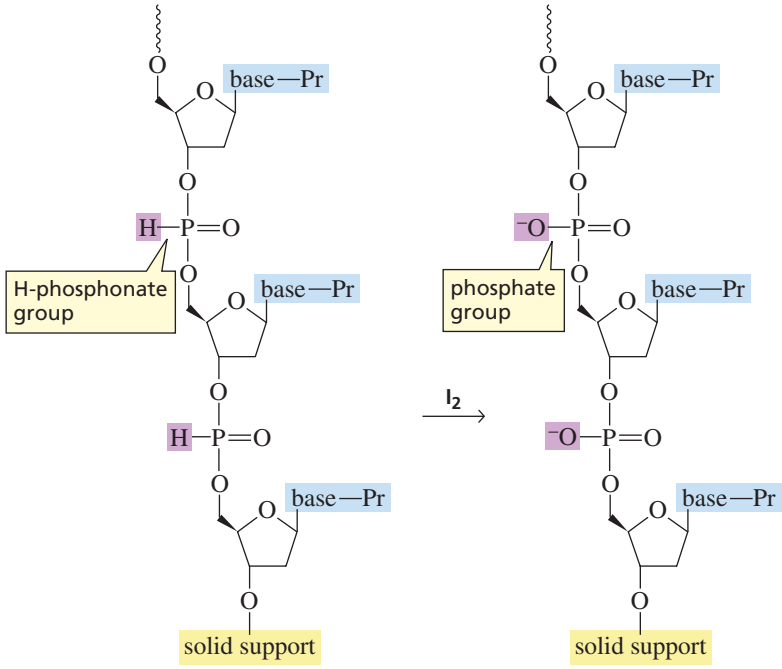
Currently, automated DNA synthesizers can synthesize nucleic acids containing as many as 130 nucleotides in acceptable yields, adding one nucleotide every 2 to 3 minutes. Longer nucleic acids can be prepared by splicing together two or more individually prepared strands. To ensure a good overall yield of oligonucleotide, the addition of each monomer must occur in high yield (>98%). This can be accomplished by using a large excess of monomer. This, however, makes oligonucleotide synthesis very expensive because the unreacted monomers are wasted.

H-Phosphonate Monomers

A second method for synthesizing oligonucleotides with specific base sequences, using H-phosphonate monomers, has the advantage over the phosphoramidite method in that the monomers are easier to handle and phosphate protecting groups

are not needed. However, the yields are not as good. The H-phosphonate monomers are activated by reacting with an acyl chloride, which converts them into phosphoanhydrides. The 5′-OH group of the nucleoside attached to the solid support reacts with the phosphoanhydride, forming a dimer. The DMTr protecting group is removed with acid under mild conditions, and a second activated monomer is added. Monomers are added one at a time in this way until the strand is complete. Oxidation with I_2 (or *tert*-butyl hydroperoxide) converts the H-phosphonate groups to phosphate groups. The base-protecting groups are removed by aqueous ammonia as in the phosphoramidite method.





Notice that in the phosphoramidite method an oxidation is done each time a monomer is added, whereas in the H-phosphonate method a single oxidation is carried out after the entire strand has been synthesized.

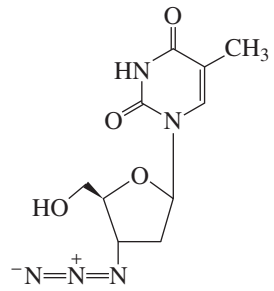
PROBLEM 27

Propose a mechanism for removal of the DMTr protecting group by treatment with acid.

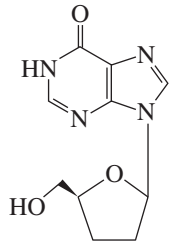
27.17 Rational Drug Design

Certain diseases such as acquired immunodeficiency syndrome (AIDS) and herpes are caused by **retroviruses**. The genetic information of a retrovirus is contained in its RNA. The retrovirus uses the sequence of bases in RNA as a template to synthesize DNA. It is called a retrovirus because its genetic information flows from RNA to DNA instead of the more typical flow from DNA to RNA.

Drugs that interfere with the synthesis of DNA by retroviruses have been designed and developed. If the retrovirus cannot synthesize DNA, it cannot take over the genetic machinery of the cell to produce more retroviral RNA and retroviral proteins. Designing drugs with particular structures to achieve specific purposes is called **rational drug design**. AZT is perhaps the best known of the drugs designed to interfere with retroviral DNA synthesis. AZT is taken up by the T lymphocytes, cells that are particularly susceptible to human immunodeficiency virus (HIV), the retrovirus that causes AIDS. Once in the cell, AZT is converted to AZT-triphosphate. The retroviral enzyme (reverse transcriptase) that catalyzes DNA formation from RNA binds AZT-triphosphate more tightly than it binds dTTP. Therefore, AZT rather than T is added to the growing DNA chain. Because AZT does not have a 3′-OH group, no additional nucleotides can be added to the chain and DNA synthesis comes to a sudden halt. Fortunately, the concentration of AZT required to affect reverse transcription is too low to affect most cellular DNA replication. A newer drug, 2′,3′-dideoxyinosine (ddI), has a similar mechanism of action. Rational drug design is further discussed in Chapter 30.



3′-azido-2′-deoxythymidine
AZT



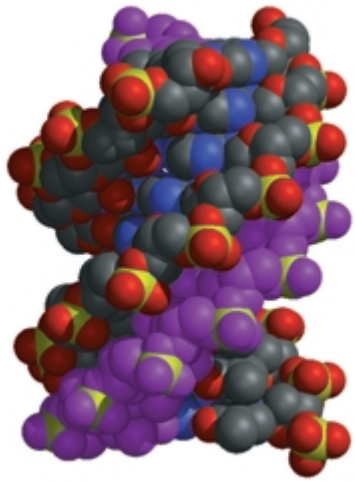
2′,3′-dideoxyinosine
ddI

3-D Molecule:
AZT

Because viruses, bacteria, and cancer cells all require DNA to grow and reproduce, chemists are trying to design compounds that will bind to the DNA of invading organisms and interfere with their reproduction. Chemists are also attempting to design polymers that will bind to specific sequences of human DNA. Such compounds could disrupt the expression of a gene (interfere with its transcription into RNA). For example, there is hope that compounds can be designed that will interfere with the expression of genes that contribute to the development of cancer. Polymers that bind to DNA are called *antigene agents*; those that bind to mRNA are called *antisense agents*.

For a compound to target a particular gene, the compound must be able to recognize a specific sequence of 15 to 20 bases. A sequence that long might occur only once in the human genome, so the compound would be specific for a particular site on DNA. In contrast, if the compound recognizes a sequence of only six bases, it could affect the human genome at more than a million locations because that sequence could occur once in every 2000 bases. However, since only 10% of the genes are expressed in most cells, a compound that recognizes a specific sequence of 10 to 12 bases may confer a gene-specific effect.

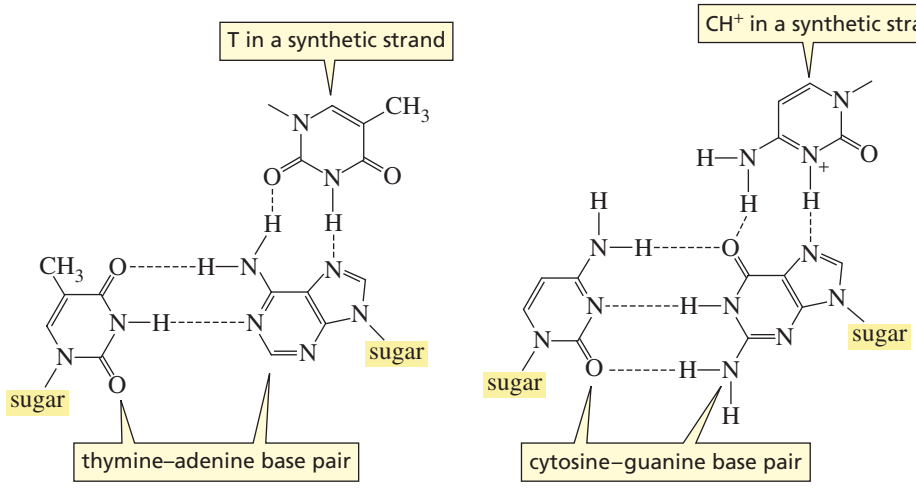
One approach to **site-specific recognition** uses a synthetic strand of oligonucleotides. When a strand of oligonucleotides is added to natural double-stranded DNA, the strand wraps around the DNA, forming a triple helix (Figure 27.19). The hope is that if the DNA sequence of a particular gene is known, a deoxyribonucleotide can be synthesized that will bind to that gene.



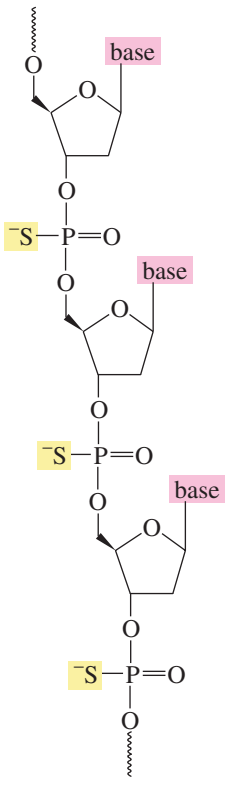
▲ Figure 27.19
A triple helix. A synthetic strand of oligonucleotides (purple atoms) is wrapped around double-stranded DNA.

3-D Molecule:
Triple helix

A triple helix is formed through *Hoogsteen base pairing* between the existing base pairs in DNA and bases in the third synthetic strand. In **Hoogsteen base pairing**, a T in the synthetic strand binds to an A of an AT base pair, and a protonated cytosine in the synthetic strand binds to a G of a GC base pair (Figure 27.20). Thus, oligonucleotides can be prepared with sequences that will base-pair to the sense strand of



◀ Figure 27.20
Hoogsteen base pairing: A T in a synthetic strand of oligonucleotides binds to the A of an A—T base pair in double-stranded DNA; a protonated C ($^{+}$CH) in the synthetic strand binds to the G of a G—C base pair in DNA.



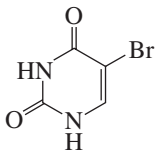
▲ **Figure 27.21**
A synthetic oligonucleotide with negatively charged sulfurs in place of negatively charged oxygens.

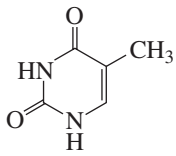
DNA at the desired location. Several methods are being investigated that will cut out the piece of targeted DNA after the synthetic strand binds to the double helix.

A problem with using synthetic oligonucleotides to target DNA is that the synthetic strands are susceptible to enzymes, such as restriction endonucleases, that catalyze the hydrolysis of DNA. Consequently, other polymers are being designed that will recognize specific DNA sequences but will not be enzymatically hydrolyzed. A compound that has shown some promise is a polymer of phosphorothioates. The polymer differs from DNA in that the negatively charged oxygen bonded to the phosphorus is replaced by a negatively charged sulfur (Figure 27.21). The polymer binds to both DNA and RNA with complementary base pairing. Oligonucleotide phosphorothioates consisting of various lengths of deoxycytidine residues have recently been found to be effective in protecting T cells from the HIV virus.

PROBLEM 28

5-Bromouracil, a highly mutagenic compound, is used in cancer chemotherapy. When administered to a patient, it is converted to the triphosphate and incorporated into DNA in place of thymine, which it resembles sterically. Why does it cause mutations? (*Hint*: The bromo substituent increases the stability of the enol tautomer.)


5-bromouracil


thymine

Summary

AU: delete prime here?

There are two types of nucleic acids—**deoxyribonucleic acid (DNA)** and **ribonucleic acid (RNA)**. DNA encodes an organism's hereditary information and controls the growth and division of cells. In most organisms the genetic information stored in DNA is **transcribed** into RNA. This information can then be **translated** for the synthesis of all the proteins needed for cellular structure and function.

ATP is a cell's most important source of chemical energy; ATP provides a reaction pathway involving a good leaving group for a reaction that would not otherwise occur because of a poor leaving group. This occurs by way of a **phosphoryl transfer reaction** in which a phosphate-containing group of ATP is transferred to a nucleophile as a result of breaking a **phosphoanhydride bond**. The reaction involves one of three intermediates—an **acyl phosphate**, an **acyl pyrophosphate**, or an **acyl adenylate**. Cleavage of a phosphoanhydride bond is highly exergonic because of electrostatic repulsions, solvation, and electron delocalization

A **nucleoside** contains a base bonded to D-ribose or to 2-deoxy-D-ribose. A **nucleotide** is a nucleoside with either the 5′- or the 3′-OH group bonded to phosphoric acid by an ester linkage. **Nucleic acids** are composed of long strands of nucleotide subunits linked by phosphodiester bonds. These linkages join the 3′-OH group of one nucleotide to the 5′-OH group of the next nucleotide. A **dinucleotide** contains two nucleotide subunits, an **oligonucleotide** contains three to ten subunits, and a **polynucleotide** contains many subunits. DNA contains 2′-deoxy-D-ribose while RNA contains D-ribose. The difference in the sugars causes DNA to be stable and RNA to be easily cleaved.

The **primary structure** of a nucleic acid is the sequence of bases in its strand. DNA contains **A**, **G**, **C**, and **T**; RNA contains **A**, **G**, **C**, and **U**. RNA has thymine instead of uracil to prevent mutations caused by imine hydrolysis of C to form U. DNA is a double-stranded helix with a major and a minor groove; the strands run in opposite directions and are twisted into a helix. The bases are confined to the inside of the helix and the sugar and phosphate groups are on the outside. The strands are held together by hydrogen bonds between bases of opposing strands as well as by **stacking interactions**—van der Waals attractions between adjacent bases on the same strand. The two strands—one is called a **sense strand** and the other a **template strand**—are complementary: **A** pairs with **T**, and **G** pairs with **C**. DNA is synthesized in the 5′ ⟶ 3′ direction by a process called **semiconservative replication**.

The sequence of bases in DNA provides the blueprint for the synthesis (**transcription**) of RNA. RNA is synthesized in the 5′ ⟶ 3′ direction by transcribing the DNA

template strand in the 3′ ⟶ 5′ direction. There are three kinds of RNA: messenger RNA, ribosomal RNA, and transfer RNA. Protein synthesis (**translation**) takes place from the N-terminal end to the C-terminal end by reading the bases along the mRNA strand in the 5′ ⟶ 3′ direction. Each three-base sequence—**a codon**—specifies the particular amino to be incorporated into a protein. A tRNA carries the amino acid bound as an ester to its 3′-terminal position. The codons and the amino acids they specify are known as the **genetic code**.

Restriction endonucleases cleave DNA at specific palindromes forming **restriction fragments**. The **dideoxy method** is the preferred method to determine the sequence of bases in the restriction fragments. Oligonucleotides with specific base sequences can be synthesized using phosphoramidite monomers or H-phosphonate monomers. Polymers that bind to DNA are called antigene agents; those that bind to RNA are called antisense agents.

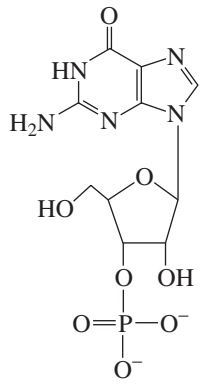
Key Terms

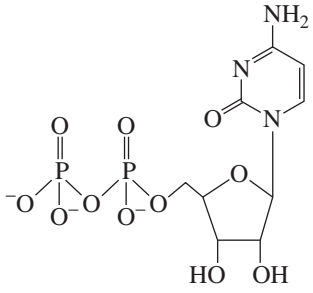
AU: Please check Key Term on page 036

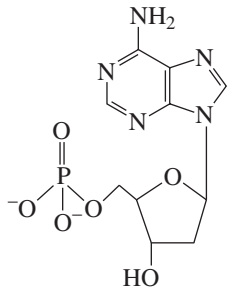
acyl adenylate (p. 040)
acyl phosphate (p. 039)
acyl pyrophosphate (p. 040)
adenosine triphosphate (ATP) (p. 036)
anticodon (p. 052)
antisense strand (p. 050)
autoradiograph (p. 061)
base (p. 032)
codon (p. 054)
deamination (p. 058)
deoxyribonucleic acid (DNA) (p. 032)
deoxyribonucleotide (p. 035)
dideoxy method (p. 060)
dinucleotide (p. 044)
double helix (p. 047)
eukaryotic organism (p. 051)
exon (p. 051)
gene (p. 050)
gene therapy (p. 061)
genetic code (p. 054)
high-energy bond (p. 041)
Hoogsteen base pairing (p. 067)
human genome (p. 050)
informational strand (p. 050)
intron (p. 051)
major groove (p. 048)
messenger RNA (mRNA) (p. 051)
minor groove (p. 048)
nucleic acid (p. 032)
nucleoside (p. 034)
nucleotide (p. 035)
oligonucleotide (p. 044)
phosphoanhydride bond (p. 037)
phosphodiester (p. 032)
phosphoryl transfer reaction (p. 038)
polynucleotide (p. 044)
primary structure (p. 044)
prokaryotic organism (p. 051)
promoter site (p. 050)
rational drug design (p. 066)
replication (p. 050)
replication fork (p. 050)
restriction endonuclease (p. 059)
restriction fragment (p. 059)
retrovirus (p. 066)
ribonucleic acid (RNA) (p. 032)
ribonucleotide (p. 035)
ribosomal RNA (rRNA) (p. 051)
ribosome (p. 051)
ribozyme (p. 051)
RNA splicing (p. 051)
sedimentation constant (p. 052)
semiconservative replication (p. 050)
sense strand (p. 050)
site-specific recognition (p. 067)
stacking interactions (p. 047)
stop codon (p. 054)
template strand (p. 050)
transcription (p. 050)
transfer RNA (tRNA) (p. 051)
translation (p. 054)

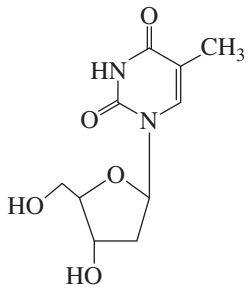
Problems

29. Name the following compounds:

a. 

b. 

c. 

d. 

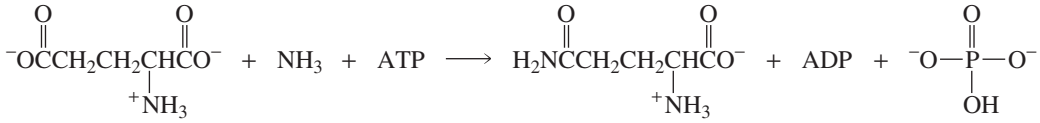
30. What nonapeptide is coded for by the following piece of mRNA?

5′—AAA—GUU—GGC—UAC—CCC—GGA—AUG—GUG—GUC—3′

31. What would be the base sequence of the segment of DNA that is responsible for the biosynthesis of the following hexapeptide?

Gly-Ser-Arg-Val-His-Glu

32. Propose a mechanism for the following reaction:



33. Match the codon with the anticodon:

Codon	Anticodon
AAA	ACC
GCA	CCU
CUU	UUU
AGG	AGG
CCU	UGA
GGU	AAG
UCA	GUC
GAC	UGC

34. Using the single-letter abbreviations for the amino acids in Table 23.1, write the sequence of amino acids in a polypeptide represented by the first four letters in your first name. Do not use any letter twice. (Because not all letters are assigned to amino acids, you might have to use one or two letters in your last name.) Write the sequence of bases in mRNA that would result in the synthesis of that polypeptide. Write the sequence of bases in the sense strand of DNA that would result in formation of the appropriate mRNA.

35. Which of the following pairs of dinucleotides are present in equal amounts in DNA?
    a. CC and GG
    b. CG and GT
    c. CA and TG
    d. CG and AT
    e. GT and CA
    f. TA and AT

36. Why is the codon a triplet rather than a doublet or a quartet?

37. RNAase, the enzyme that catalyzes the hydrolysis of RNA, has two catalytically active histidine residues at its active site. One of the histidine residues is catalytically active in its acidic form and the other is catalytically active in its basic form. Propose a mechanism for RNAase.

38. The amino acid sequences of peptide fragments obtained from a normal protein and from the same protein synthesized by a defective gene were compared. They were found to differ in only one peptide fragment. The primary sequences of the fragments are shown here.

    Normal: Gln-Tyr-Gly-Thr-Arg-Tyr-Val
    Mutant: Gln-Ser-Glu-Pro-Gly-Thr

    a. What is the defect in DNA?
    b. It was later determined that the normal peptide fragment is an octapeptide with a C-terminal Val-Leu. What is the C-terminal amino acid of the mutant peptide?

39. Whether the mechanism requiring activation of a carboxylate ion by ATP involves attack of the carboxylate ion on the α-phosphorus or the β-phosphorus of ATP cannot be determined from the reaction products because AMP and pyrophosphate are obtained as products in both mechanisms. The mechanisms, however, can be distinguished by a labeling experiment in which the enzyme, the carboxylate ion, ATP, and radioactively labeled pyrophosphate are incubated, and the ATP is isolated. If the isolated ATP is radioactive, the mechanism involves attack on the α-phosphorus. If it is not radioactive, the mechanism involves attack on the β-phosphorus. Explain these conclusions.

40. What would be the results of the experiment in Problem 39 if radioactive AMP were added to the incubation mixture instead of radioactive pyrophosphate?

41. Which cytosine in the following sense strand of DNA could cause the most damage to the organism if it were deaminated?

    5′—A—T—G—T—C—G—C—T—A—A—T—C—3′

X-ref 42. Sodium nitrite, a common food preservative (p. 000), is capable of causing mutations in an acidic environment by converting cytosines to uracils. Explain how this occurs.

43. The first amino acid incorporated into a polypeptide chain during its biosynthesis in prokaryotes is *N*-formylmethionine. Explain the purpose of the formyl group.

PART EIGHT

Special Topics in Organic Chemistry

In previous chapters, the polymers synthesized by biological systems—proteins, carbohydrates, and nucleic acids—were discussed.

Chapter 28 discusses polymers synthesized by chemists. These synthetic polymers have physical properties that make them useful in everyday life, in applications such as fabrics, bottles, food wrap, automobile parts, and compact discs.

Chapter 29 discusses pericyclic reactions—reactions that occur as a result of a cyclic reorganization of electrons. In this chapter, you will learn how the conservation of orbital symmetry theory explains the relationships among reactant, product, and reaction conditions in a pericyclic reaction.

Chapter 30 introduces you to medicinal chemistry. Here you will see how many of our commonly used drugs were discovered, and you will learn about some of the techniques used to develop new drugs.